INAPPROPRIATE EXPRESSIONS RECOGNITION USING BOOTSTRAPPING AS SEMI-SUPERVISED LEARNING

A Thesis
Presented to the Faculty of the
College of Computer and Information Sciences
Polytechnic University of the Philippines

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Science in Computer Science

Dapitan, Joshua S.
Lasala, Anjanette R.

March 2016

SIGNATURE PAGE

The thesis **"Inappropriate Expressions Recognition using Bootstrapping as Semi-Supervised Learning"** submitted and presented by Joshua S. Dapitan and Anjanette R. Lasala in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science has been

Recommended for Approval and Acceptance: Date:

______________________________ ____________
Dr. Iluminada Vivien Domingo, MBA, DBA
Faculty-In-charge

______________________________ ____________
Prof. Ria A. Sagum
Thesis/Technical Advisor

Approved by the Thesis Defense Panel:

______________________________ ____________
<NAME>

______________________________ ____________
<NAME>

______________________________ ____________
<NAME>

Accepted for the Department of Computer Science

______________________________ ____________
Michael B. dela Fuente, MSGITS
Department Chairperson

Accepted for the College of Computer and Information Sciences

______________________________ ____________
Gisela May Albano
Dean of the College

ABSTRACT

Inappropriate expressions recognition is a task of recognizing words whose usage is in an inappropriate context, which is the words used in offensive sense or sexually explicit sense. In this study, the researchers developed a prototype inappropriate expressions recognition system that analyzes sentences in a phrase-level orientation to determine if the word usage is inappropriate or not. Our approach is by using Bootstrapping, Naïve Bayesian Classification, N-Gram Language Modelling, Bag of Words Model, and Hidden Markov Modelling. We used Stanford NLP tools for some parts of the pre-processing. Our dictionaries that serve as a basis of defining inappropriateness for the prototype is WordNet 2.1 and UrbanDictionary, which definitions are extracted by Python's BeautifulSoup and Requests API. We used randomly chosen comments from websites such as 9gag and YouTube as input for the testing and experiment of the system. After the experiment is executed, considering 500 comments, the following rates were computed based from the gathered data. Recall – 66.84%, Precision – 73.12%, Specificity – 96.70% and F-Measure – 69.84%. We were able to conclude that using the definitions as the basis for inappropriate expressions recognition is possible to give possible results and Inappropriate expressions can be modeled despite the noise produced by the informal definitions of UrbanDictionary.

ACKNOWLEDGMENT

This thesis becomes a reality with the kind of support and help of many individuals. We would like to extend our sincere thanks to all of them.

To our parents, Mr. and Mrs. Dapitan and Lasala, for the encouragement, heartening support and understanding they have shown in every task we do.

To our classmates and friends, for extending their encouraging support in the process of constructing this paper even beyond their busy schedules. All the positivity for us as we continue our walk in this difficult yet fulfilling journey.

To Dr. Vivien Domingo, for her constructive criticisms that guided us in appropriately painting the words that this paper contains.

To our Thesis Adviser, Prof. Ria Sagum, for imparting her knowledge and expertise in this study and for giving helpful comments and suggestions for the progress of the researchers' paper.

Foremost, we want to offer this endeavor to our GOD Almighty for the wisdom he bestowed upon us, the strength, peace of mind and good health in order to finish this research.

The Researchers

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

THE PROBLEM AND ITS BACKGROUND

This chapter discusses the introduction, background of the study, conceptual framework, and statement of the problem, scope and limitation, significance of the study, hypothesis and definition of terms. It also states the objectives and target question that needs to be answered after conducting the study.

1.1 Introduction

Inappropriate Expressions is one of the problems in a behavioral sense. Inappropriate Expressions mostly causes problems in literary management like cyber bullying, and exposure of children to other textual data that may cause other interests like crime, sex, etc.

The solution is to analyze the inappropriate expressions and model the inappropriate language. The problem is, due to the inherent ambiguity of the language, there is a hard time to recognize the real meaning if whether the expression gets inappropriate, making it harder to be recognized. For example, the word screw may mean the actual screw material, or a slang term for sexual intercourse, which is in accordance to the definition of WordNet of the said example. There are also implemented solutions such as word filters which is by detection of the words, which affects the recognition because of the disregarding of the context.

The solution to be used in this study is a machine learning methodology, which is bootstrapping, which is designed to model the Inappropriate language, in which embodies Inappropriate Expressions. With this solution, users may find inappropriate expressions in textual data, which can be used as a tool for prevention of the exposure of the inappropriate expressions to those who are not concerned. This also models the sentence-level context analysis to identify the inappropriateness of an expression with the use of Lexical Syntactic Features and Grammar Relations as a support to the said computational model that is used to solve the problem of modeling the inappropriate language.

1.2 Background of the Study

In many ways the Internet is like a gigantic library; both have content to teach and entertain. And similar to the content in a library, not all Internet content is appropriate for children. Libraries create children's and young adults' sections in order to help youths (and their parents) identify which materials are appropriate for them. On the Internet, however, all of the content may be equally accessible; websites about ponies and websites featuring pornography are both a click away. 87% of children go online at home and it is possible that they can read some inappropriate expressions on the internet. These inappropriate expressions could be expressions containing swearing, unmoderated chatrooms where there's no one supervising the conversation and barring unsuitable comments and sexual explicit comments [Inappropriate Content, 2015].

These inappropriate expressions become very interesting in the field of NLP Community. Fortunately, some researches like CAW 2.0 build a system that detects cyber bullying on the internet. It is determined that a baseline text mining system (using bag of words approach) was significantly improved by including sentiment and contextual features. Even with the combined model, a support vector machine learner could only produce a recall level of 61.9% [Hong et. al., 2009]. And another research that was built was profanity related offensive content in twitter. The researchers introduced an approach that exploits linguistic regularities in profane language via statistical topic modeling on a huge Twitter corpus, and detects offensive tweets using these automatically generated features. The keyword matching technique has been shown to perform very well in the literature and achieved a TP of 69.7% with an FP of 3.77% in our experiment. While keeping the FP on the same level as the baseline, our approach had a TP of 75.1% over 4029 testing tweets using Logistic Regression, a significant 5.4% improvement over the baseline [Fan et.al, n.d.]. Therefore, one of the problems that need to be solve in this study is the accuracy of the system because some of the researches related to the study has a low rate of accuracy.

In Sentiment classification, it aims to predict the sentiment polarity of a text [Lee et al., 2002] and plays a critical role in many Natural Language Processing (NLP) applications [Cheng et al., 2005; Cardie et al., 2005; Cui et al., 2006; Balahur et al., 2009]. Although supervised learning methods have been widely employed and proven effective in sentiment classification in the literature [Lee et al., 2002], they normally depend on a large amount of labeled data, which usually involves high cost in labor and time. To overcome this problem, various semi-supervised learning methods are proposed to effectively utilize a small scale of labeled data along with a larger amount of unlabeled data.

One specific open problems in semi-supervised 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 non-trivial way and still get an efficient algorithm for this class [Balcan and Blum, n.d.].

However, all the existing semi-supervised learning methods assume the balance between negative and positive samples in both the labeled and unlabeled data, and none of them consider a more common case where the class distribution is imbalanced, i.e., the number of positive samples is quite different from that of negative samples in both the labeled and unlabeled data. For clarity, the class with more samples is referred as the majority class (MA) and the other class with fewer samples is referred as the minority class (MI). In fact, semi-supervised learning on imbalanced classification is rather challenging: at least, there exist two basic issues to be solved. On the one hand, imbalanced classification requires a specifically-designed classification algorithm. Trained on the imbalanced labeled data, most classification algorithms tend to predict test samples as the majority class and may ignore the minority class. Although many methods, such as re-sampling [Bowyer et al., 2002], one-class classification [Duin and Juszczak, 2003], and cost-sensitive learning [Liu and Zhou, 2006], have been proposed to solve this issue, it is still unclear as to which method is more suitable to handle the imbalanced problem in sentiment classification and whether the method is extendable to semi-supervised learning.

The solution to be used in this study is a machine learning methodology, which is bootstrapping, which is designed to model the Inappropriate language, in which embodies Inappropriate Expressions. With this solution, users may find inappropriate expressions in textual data, which can be used as a tool for prevention of the exposure of the inappropriate expressions to those who are not concerned.

1.3 Theoretical Framework

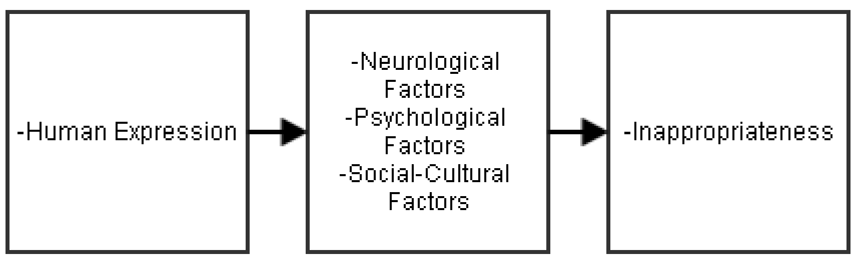


Figure 1 – Neuro-Psycho-Social Theory

The theory, which is Neuro-Psycho-Social Theory of Speech. There are "Rules" that governs the human's expression of Inappropriateness, Offensiveness, and Humor. The catalysts of these expressions are the Neurological Factors, Psychological Factors, and Social-Cultural Factors. [Jay, 2009]

1.3 Conceptual Framework

1.4.1 Conceptual Framework of the System

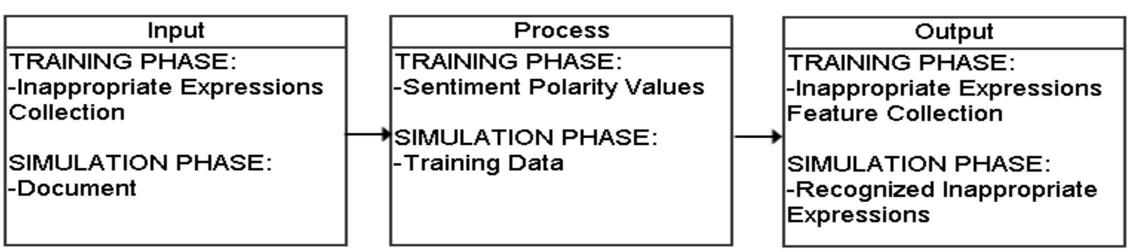


Figure 2 – Conceptual Framework of the System

The concept was to create a model in which there would be a. machine that learns inappropriate expressions and recognize it. The process of training would be by the utilization of a sentiment corpus to represent the catalysts of the Inappropriate expression, which contains the sentiment polarity values, which would affect the threshold values for the inappropriateness and the features that would have been collected in the inappropriate expressions feature collection, in which leads to affect the performance of the recognition of inappropriate expressions during the simulation.

1.4.2 Conceptual Framework of the Study

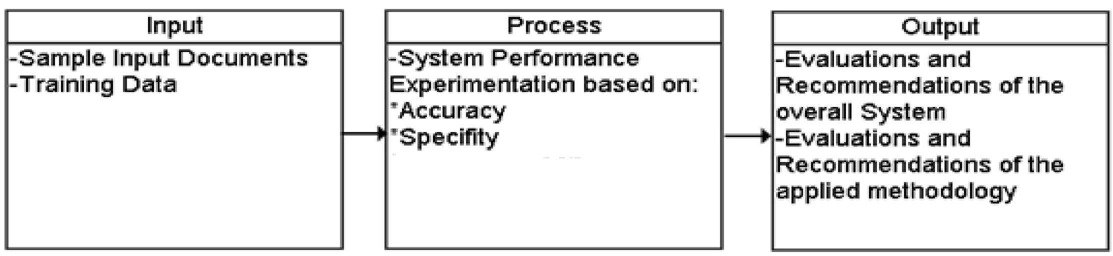


Figure 3 – Conceptual Framework of the Study

The concept was to study a model in which there would be a. machine that learns inappropriate expressions and recognize it. There would be different training sets and different sample documents. Then, the results would be experimented via Experiment Paper. The output would be the evaluations and recommendations of the system and the approach to solve the problem about recognizing inappropriate expressions.

1.4 Statement of the Problem

The study was designed to create and evaluate the Model that recognizes Inappropriate Expressions from a document. In addition to this, the researchers sought an answer to this problem:

1. What is the performance analysis of the model in terms of:

    1.1 Recognition of Inappropriate expressions. (Accuracy of the Model)

    1.2 Recognition of Appropriate expressions. (Specificity of the Model)

1.5 Scope and Limitations

1.5.1 Scope and Limitation of the System

The Inappropriate Expressions Recognizer focused on the recognition of Contextually Inappropriate Expressions in English comments. The user may input comments in a given textbox. It scans for Inappropriate Expressions and their usages in the input comment to verify the Inappropriate Expressions' inappropriateness. Proper Capitalization is advised and symbols _ and / is restricted. The Inappropriate Expressions that are accommodated for are limited to one word inappropriate expressions. Multiple word inappropriate expressions are unaccommodated due to the ambiguous nature of these. The domain for the recognition of inappropriateness is Offensive Language and Sexually Explicit Language.

1.5.2 Scope and Limitations of the Study

Since "Inappropriate Expressions Recognition using bootstrapping as Semi-Supervised Learning" was an experimental research, the system went through several testing and evaluation to test the accuracy for different types of inputs.

1.6 Significance of the Study

Inappropriate Expressions Recognition using Bootstrapping as Semi-Supervised Learning focuses on recognizing inappropriate expressions on a document. The system will benefit the following:

Students – these are people who have a whole society around technology because most of them routinely use chat and email to communicate with each other. The system will prevent receiving possible inappropriate expressions through the internet.

Parents – these are the people who are worrying so much about what their children do on the internet because some articles found on google are not appropriate for their child and contains inappropriate expressions like offensive and sexually explicit words.

Editors- these are the people who are worrying so much about the works that will be published. With this model, there will be recognition of inappropriate expressions in the document that may help the editorial judge the work's level of inappropriateness.

Moderators- these are the people who manage the behavior of users in an online community. With this model, the moderators can have ease in reading the activities that has to deal with inappropriate expressions.

Natural Language Processing Researchers (NLP) – These are the experts in Natural Language processing. The system will provide avenues for further improvement of the said topic.

Future Researches – These are the people who will conduct future research to improve the existing studies about Sentiment Analysis and Inappropriate Expressions. They can add additional functions to the system.

1.7 Definition of Terms

Bootstrapping Algorithm– is a method for deriving robust estimates of standard errors and confidence intervals for estimates such as the mean, median, proportion, odds ratio, correlation coefficient or regression coefficient.

F-Measure - is the weighted average of the values of the Precision and Recall.

Inappropriate Expressions – is something that is not within the bounds of what is considered appropriate or socially acceptable.

Naive Bayes - methods are a set of supervised learning algorithms based on applying Bayes' theorem with the "naive" assumption of independence between every pair of features.

Named Entity Recognizer - is a piece of software that reads text in some language and assigns a classification of proper nouns such as person, location, organization, etc.

Natural Language Processing – is a field of computer science, artificial intelligence and computational linguistics concerned with the interactions between computers and human languages.

N-Gram - is a contiguous sequence of words from a given sequence of text.

Offensive Language – the term that is applied to hurtful, derogatory or obscene comments made by one person to another person.

Part-of-Speech Tagger - is a piece of software that reads text in some language and assigns parts of speech to each word (and other token), such as noun, verb, adjective, etc., although generally computational applications use more fine-grained POS tags like 'noun-plural'.

Precision - Percentage of identified expressions that are inappropriate.

Recall - Percentage of inappropriate expressions correctly identified.

Semi-supervised Learning – is a class of supervised learning tasks and techniques that also make use of unlabeled data for training – typically a small amount of labeled data with a large amount of unlabeled data.

Sentiment Analysis – refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials.

SentiWordNet - is a lexical resource for opinion mining. SentiWordNet assigns to each synset of WordNet three sentiment scores: positivity, negativity, objectivity.

Sexually explicit – presents sexual content without deliberately obscuring or censoring it.

Specificity - is the rate of the results without the condition, which has a negative test result.

Urban Dictionary – is comprised of submissions from everyday people and regulated by volunteer editors, who are given an extremely small set of rules to maintain quality control.

WordNet - is a lexical database for the English language. It groups English words into sets of synonyms called synsets, provides short definitions and usage examples, and records a number of relations among these synonym sets or their members.

CHAPTER II

REVIEW OF RELATED LITERATURE AND STUDIES

This chapter includes review of related literature and studies. It also includes the synthesis of the study. This chapter provided relevant information and researchers regarding the study that gave additional knowledge to the researchers.

2.1 Review of Related Literature

Swear words (or taboo words, as he calls them) can include sexual references (*fuck*), those that are profane or blasphemous (*goddamn*), scatological or disgusting objects (*shit*), animal names (*pig, ass*), ethnic/racial/gender slurs (*fag*), ancestral allusions (*bastard*), substandard vulgar terms and offensive slang. Taboo words can be mildly offensive to extremely offensive, and people will often use a milder euphemism to replace a swear word when in mixed (or unknown) company [Jay, 2009].

Taboo words can be used for a variety of reasons, including to achieve a specific reaction from others. Swearing injects a direct, succinct emotional component into the discussion, usually in order to express frustration, anger or surprise (up to two-thirds of our swearing is for just such expressions). These insulting swears can be name calling or wishing someone harm, so it’s no wonder they are often a defining feature of hate speech, verbal abuse, sexual harassment and obscene phone calls [Grohol, 2009].

Cursing appears as soon as children hear curse words, as early as one year of age [Jay, 2009]. Children's cursing emerges in a predictable fashion. Early cursing and name calling are based on references to scatology and perceived differences about others (e.g., not-eater, four-eyes). In adolescence, cursing becomes more abstract and socially based. Gender differences in cursing emerge as soon as children attend school: Boys curse more and use more words and use more offensive words than do girls. Cursing reaches a peak in adolescence but continues into old age, persisting through senile decline and dementia.

Children learn that curse words are associated with emotion states through classical conditioning, the repeated pairing of words (e.g., damn!) with emotional events. Curse words effectively replace infantile expressions of anger such as biting and screaming [Goodenough, 1931]. Children associate curse words with all emotion states (e.g., joy, surprise, fear); they learn to express emotions through words, and they learn to perceive others' emotional states through the emotional speech they observe. Children learn that curse words intensify emotions in a manner that noncurse words cannot achieve.

Language learning and cursing depend on one's social, emotional, and cognitive reasoning abilities. As children become more cognitive1y sophisticated, their emotional language, name calling, and sexual references shift to match their higher mental functioning. Many uses of curse words occur at an automatic or reflexive level in the form of response cries and epithets. Eventually, the semantic and syntactic rules for cursing are acquired, allowing children to use curse words appropriately in propositional statements [Jay, 2009]. These propositional statements are primarily used to express emotions (connotation), but curse words also function to make references about the world (denotation).

Cursing is never chaotic, meaningless, or random behavior – cursing is seen as purposeful and rule-governed. The rule of NPS Theory is to generate a likelihood “rules” that underlie concepts of appropriateness, offensiveness and humor. Native speakers acquire cursing rules as they learn language. Discovering and testing these cursing rules is meant to give the theory predictive power. The more accurately the NPS Theory can predict acts of cursing, the more valid is our understanding of cursing [Jay, 2009].

Human sexuality is a critical aspect of emotional language in general and of cursing in particular because sexuality is one of the most tabooed aspects of human existence. The language of sexuality is intimately connected to one's emotional life, one's sexual orientation, and one's cursing habits or style. Human sexuality becomes represented in two ways: The sexual body is represented as a materiality, and a set of sexual ideas or sexual language is developed about that materiality [Jay, 2009].

Children learn sexual terminology through interactions with peers and adults. Parents express their sexual values, fears, and anxieties to children when they inhibit or punish sexual references. Punishment and avoidance of sex terms teach the child that sexual words are powerful and that sexuality itself is powerful. Parents with high sexual anxiety are likely to transfer their anxiety to their children, who learn that both sex talk and sex are to be avoided. This learning takes place through the repression of sex talk itself through a course of negations and omissions. Through the acquisition of sexual terminology and the conditioned fears and pleasures regarding sexuality, the child develops a level of comfort with sexuality. This sexual identity will influence how a speaker uses words for sex acts, body parts, and gender related insults with other people [Jay, 2009].

As children develop linguistically and sexually, their conversations about sexuality become more highly dependent on who is listening. Both adolescent and adult sexual conversations clearly depend on intimacy, sexual identity, and formality [Wells, 1990]. Most adults can talk about sexuality with lovers or with others who share similar sexual preferences. But almost all young adults have trouble talking about sex with their parents (those who avoided sex talk in the first place) and in mixed gender crowds.

Most people talk about sex by using vulgar terms and sexual slang [Jay, 2009]. Clinical terms are reserved for polite situations. Some sex acts are so taboo (e.g., oral sex) that no acceptable term can be used in polite company. Euphemisms and circumlocutions are commonly used in order to talk about sex and taboo topics. In fact, married couples, cohabiting couples, and sexually active partners create personal idioms and idiosyncratic terminology to use in intimate situations [Cornog, 1986]. We see in the following chapters in Part ill that the use of sexual language is very important to speakers, revealing their personality traits, attitudes about sexuality, and parental influences. Both the physical acts of sex and one's sexual identity are expressed and experienced through language choices.

Inappropriate Expression is something that is not within the bounds of what is considered appropriate or socially acceptable [YourDictionary, n.d.]. Inappropriate expression includes information that upset us or information that might lead or tempt us into unlawful or dangerous behavior. This could be content containing swearing, unmoderated chatrooms where there’s no one supervising the conversation and barring unsuitable comments and sexual explicit comments [Inappropriate Content, n.d.]. Profanity is an offensive word or inappropriate 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 a debasement 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 behavior showing similar disrespect [Longman Dictionary of Contemporary English, 2014].

The first cognitive system was Watson, which debuted in a televised Jeopardy! Challenge where it bested the show's two greatest champions. The challenge for Watson was to answer questions posed in every nuance of natural language, such as puns, synonyms and homonyms, slang, and jargon [IBM Research, n.d.].

In 2011 Watson, the IBM super computer best known for its run as a Jeopardy contestant, gained a new tool in its language arsenal: swearing [How-to-Geek, n.d.]. Researchers in charge of expanding Watson's vocabulary and ability to use language in a more nuanced and natural fashion thought it would be helpful to teach Watson slang and colloquial sayings. Essentially, they wanted to give Watson the ability to speak more like we speak to each other and less like a super computer carefully selecting an answer [How-to-Geek, n.d.].

To this end the researchers unleashed Watson on Urban Dictionary, the massive 7-million-entry Internet dictionary of slang words and phrases. On one hand the experiment was a huge success, crunching through the Urban Dictionary database radically expanded Watson's word selection and the nuance of its language use. On the other hand, Watson proved to be terrible at distinguishing when its language was and was not appropriate–shortly after it acquired its new vocabulary from Urban Dictionary it reportedly responded to a researcher's inquiry by reporting it was "bullshit" [How-to-Geek, n.d.].

In response to the changes in Watson's vocabulary researchers terminated the Urban Dictionary experiment and set up filters to help Watson refrain from swearing in the future [How-to-Geek, n.d.].

A *swear filter*, also known as a *profanity filter* or *language filter* is a software subsystem which modifies text to remove words deemed offensive by the administrator or community of an online forum. Swear filters are common in customprogrammed chat rooms and online video games, primarily MMORPGs. This is not to be confused with content filtering, which is usually built into internet browsing programs by third-party developers to filter or block specific websites or types of websites. Swear filters are usually created or implemented by the developers of the Internet service [FileSharingTalk, 2006].

A common quirk with wordfilters, often considered either comical or annoying by users, is that they often affect words that are not intended to be filtered. This is a typical problem when short words are filtered. For example, if the word "ass" is filtered, so are "assist", "classic", "assassin", and other words which contain the sequence. For example, one may see, "Do you need ***istance for playing cl***ical music?" Multiple words may be filtered if whitespace is ignored, resulting in "as suspected" becoming "****uspected". Prohibiting a phrase such as "hard on" will result in filtering innocuous statements such as "That was a hard one!" and "Sorry I was hard on you." [Sheerin, 2010].

Some words that have been filtered accidentally can become replacements for profane words. One example of this is found on the Myst forum Mystcommunity. There, the word 'manuscript' was accidentally censored for containing the word 'anus', which resulted in 'm****cript'. The word was adopted as a replacement swear and carried over when the forum moved, and many substitutes, such as " 'scripting ", are used (though mostly by the older community members) [Sheerin, 2010].

Place names may be filtered out unintentionally due to containing portions of swear words. In the early years of the internet, the British place name Penistone was often filtered out from spam and swear filters [Sheerin, 2010].

Naive Bayes has been studied extensively since the 1950s. It was introduced under a different name into the text retrieval community in the early 1960s, [Norvig and Russell et.al., 2003] and remains a popular (baseline) method for text categorization, the problem of judging documents as belonging to one category or the other (such as spam or legitimate, sports or politics, etc.) with word frequencies as the features. With appropriate preprocessing, it is competitive in this domain with more advanced methods including support vector machines [Karger et.al. 2003]. It also finds application in automatic medical diagnosis [Rish, 2001].

Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum-likelihood training can be done by evaluating a closed-form expression [Norvig and Russell, 2003], which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers.

An n-gram model is a type of probabilistic language model for predicting the next item in such a sequence in the form of a $(n - 1)$–order Markov model. N-gram models are now widely used in probability, communication theory, computational linguistics (for instance, statistical natural language processing), computational biology (for instance, biological sequence analysis), and data compression. Two benefits of n-gram models (and algorithms that use them) are simplicity and scalability – with larger n, a model can store more context with a well-understood space–time tradeoff, enabling small experiments to scale up efficiently [Brants 2006].

N-gram models are widely used in statistical natural language processing. In speech recognition, phonemes and sequences of phonemes are modeled using a n-gram distribution. For parsing, words are modeled such that each n-gram is composed of n words. For language identification, sequences of characters/graphemes (e.g., letters of the alphabet) are modeled for different languages. For sequences of characters, the 3-grams (sometimes referred to as "trigrams") that can be generated from "good morning" are "goo", "ood", "od ", "d m", " mo", "mor" and so forth (sometimes the beginning and end of a text are modeled explicitly, adding "__g", "_go", "ng_", and "g__"). For sequences of words, the trigrams that can be generated from "the dog smelled like a skunk" are "# the dog", "the dog smelled", "dog smelled like", "smelled like a", "like a skunk" and "a skunk #" [Dunning, 1994].

In practice, n-gram models have been shown to be extremely effective in modeling language data, which is a core component in modern statistical language applications [Dunning, 1994].

Most modern applications that rely on n-gram based models, such as machine translation applications, do not rely exclusively on such models; instead, they typically also incorporate Bayesian inference. Modern statistical models are typically made up of two parts, a prior distribution describing the inherent likelihood of a possible result and alikelihood function used to assess the compatibility of a possible result with observed data. When a language model is used, it is used as part of the prior distribution (e.g. to gauge the inherent "goodness" of a possible translation), and even then it is often not the only component in this distribution [Dunning, 1994].

The bag-of-words model is a simplifying representation used in natural language processing and information retrieval (IR). In this model, a text (such as a sentence or a document) is represented as the bag (multiset) of its words, disregarding grammar and even word order but keeping multiplicity. Recently, the bag-of-words model has also been used for computer vision [Sivic 2009].

The bag-of-words model is commonly used in methods of document classification, where the (frequency of) occurrence of each word is used as a feature for training a classifier. An early reference to "bag of words" in a linguistic context can be found in Zellig Harris's 1954 article on *Distributional Structure* [Harris 1954].

NER systems have been created that use linguistic grammar-based techniques as well as statistical models, i.e. machine learning. Hand-crafted grammar-based systems typically obtain better precision, but at the cost of lower recall and months of work by experienced computational linguists. Statistical NER systems typically require a large amount of manually annotated training data. Semi-Supervised approaches have been suggested to avoid part of the annotation effort [Lin 2009,Northman 2013].

Many different classifier types have been used to perform machine-learned NER, with conditional random fields being a typical choice [Finkel 2005].

Semi-supervised learning is a class of supervised learning tasks and techniques that also make use of unlabeled data for training - typically a small amount of labeled data with a large amount of unlabeled data. Semi-supervised learning falls between unsupervised learning (without any labeled training data) and supervised learning (with completely labeled training data). Many machine-learning researchers have found that unlabeled data, when used in conjunction with a small amount of labeled data, can produce considerable improvement in learning accuracy [Wikipedia, 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].

The Bootstrap algorithm, which is Semi-Supervised Learning, works by drawing many independent bootstrap samples, evaluating the corresponding bootstrap replications, and estimating the standard error of $\hat{\theta}$ by the empirical standard error, denoted by $\widehat{se}_B$, where B is the number of bootstrap samples used [Efron and Tibshirani, 1993].

With the bootstrap method, the basic sample is treated as the population and a Monte Carlo-style procedure is conducted on it. This is done by randomly drawing a large number of ‘resamples’ of size n from this original sample (of size n either) with replacement. So, although each resample will have the same number of elements as the original sample, it could include some of the original data points more than once, and some not included. Therefore, each of these resamples will randomly depart from the original sample. And because the elements in these resamples vary slightly, the statistic G*, calculated from one of these resample will take on slightly different values. The central assertion of the bootstrap method is that the relative frequency distribution of these G*’s is an estimate of the sampling distribution of G [The Original Bootstrap Method, n.d.].

A hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. A HMM can be presented as the simplest dynamic Bayesian network. The mathematics behind the HMM were developed by L. E. Baum and coworkers [Baum 1966]. It is closely related to an earlier work on the optimal nonlinear filtering problem by Ruslan L. Stratonovich, who was the first to describe the forward-backward procedure [Stratonovich 1960].

In simpler Markov models (like a Markov chain), the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but the output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore, the sequence of tokens generated by an HMM gives some information about the sequence of states. The adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; the model is still referred to as a 'hidden' Markov model even if these parameters are known exactly.

Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture, part-of-speech tagging, musical score following, partial discharges and bioinformatics [Satish 2003].

A hidden Markov model can be considered a generalization of a mixture model where the hidden variables (or latent variables), which control the mixture component to be selected for each observation, are related through a Markov process rather than independent of each other. Recently, hidden Markov models have been generalized to pairwise Markov models and triplet Markov models which allow consideration of more complex data structures and the modelling of nonstationary data [Boudaren, 2012].

2.2 Review of Related Studies

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.].

Sentiment Analysis has the potential to identify victims who pose high risk to themselves or others, and to enhance the scientific understanding of bullying overall victims usually experience negative emotions such as depression, anxiety and loneliness. In extreme cases such emotions are more violent or even suicidal. Detecting at risk individuals via sentiment analysis enables potential interventions. In addition, social scientists are interested in sentiment analysis on bullying traces to understand participants' motivations [Bellmore et.al, 2012].

The Lexical Syntactical Feature (LSF) approach from the research Detecting Offensive Language in Social Media to Protect Adolescent Online Safety is to identify offensive contents in social media, and further predict a user's potentiality to send out offensive contents. It includes two phases of offensive detection. Phase 1 aims to detect the offensiveness on the sentence level and Phase 2 derives offensiveness on the user level. In Phase 1, the researchers apply advanced text mining and natural language processing technique to derive lexical and syntactic features of each sentence. Using these features, we derive an offensive value for each sentence. In Phase 2, we further incorporate user-level features where we leverage research on authorship analysis. The system consists of pre-processing and two major components: sentence offensiveness prediction and user offensiveness estimation. During the pre-processing stage, user's conversation history is chunked into posts, and then into sentences. During sentence offensiveness prediction, each sentence's offensiveness can be derived from two features: its word's offensiveness and the context. The researchers use lexical feature to represent words' offensiveness in a sentence, and syntactic feature to represent context in a sentence. Words' offensiveness nature is measured from two lexicons. For the context, we grammatically parse sentences into dependency sets to capture all dependency types between a word and other words in the same sentence, and mark some of its related words as intensifiers. The intensifiers are effective in detecting whether offensive words are used to describe users or other offensive words. During user offensiveness estimation stage, sentence offensiveness and users' language patterns are helped to predict user's likelihood of being offensive. Experimental result shows that the LSF sentence offensiveness prediction and user offensiveness estimate algorithms outperform traditional learning based approaches in terms of precision, recall and f-score. It also achieves high processing speed for effective deployment in social media [Chen et.al, n.d.].

Very few other research teams are working on the detection of cyber bullying. A misbehavior detection task was offered by the organizers of CAW 2.0, but only one submission was received. It is determined that a baseline text mining system (using bag of words approach) was significantly improved by including sentiment and contextual features. Even with the combined model, a support vector machine learner could only produce a recall level of 61.9% [Hong et. al., 2009].

In the research Filtering Offensive Language in Online Communities using Grammatical Relations, mainly tackles the problem about how the offensive language can be removed in a user message. They analyze the offensive language in text messages posted in online communities, and propose a new automatic sentence-level filtering approach that is able to semantically remove the offensive language utilizing the grammatical relations among words. Their solution includes 3 steps. First, they analyzed the characteristics of offensive text content in user messages. Then, they introduced their filtering philosophy according to the summarized characteristics. Finally, they show how this philosophy is transformed into heuristic rules applicable in the filtering process. The overview idea of their filtering approach is shown in the presented Algorithm 1 in which the functions POS tagging ad TD generator generate Part-of-Speech tags and typed dependency relations, respectively. They also use existing NLP (Natural Language Processing) tools to implement these two functions. They also focused in the design of two other functions CreateRelTree and EstimateRelTree. In their research assume that the filtering is based on a comprehensive offensive lexicon containing all offensive words. Words do not appear in the lexicon are considered inoffensive. Experiments their dataset, comments from Youtube, show over 90% agreement in filtered results between the proposed approach and manual filtering approach [Xu and Sencun, 2010].

The researchers of Detecting Offensive Tweets via Topical Feature Discovery over a Large Scale Twitter Corpus in proposed a novel semi-supervised approach for detecting profanity-related offensive content in Twitter. They introduced an approach that exploits linguistic regularities in profane language via statistical topic modeling on a huge Twitter corpus, and detects offensive tweets using these automatically generated features. Their step by step processes are as follows: (a) Bootstrap between twitters and tweets based on a seed word set to obtain training tweets for topic model learning; (b) Topic models are learned via a generative LDA approach; (c) Tweets in a holdout testing set are processed in the same fashion as in (a); (d) Topic distributions are inferred for each testing tweet by the topic model learned in step (b); (e) Seed words are applied against each testing tweet, leading to a binary lexicon feature; (f) ML models are built and evaluated. The keyword matching technique has been shown to perform very well in the literature and achieved a TP of 69.7% with an FP of 3.77% in our experiment. While keeping the FP on the same level as the baseline, our approach had a TP of 75.1% over 4029 testing tweets using Logistic Regression, a significant 5.4% improvement over the baseline [Fan et.al, n.d.].

Sentiment classification methods can be categorized into three types: unsupervised [Turney, 2002], supervised [Pang et al., 2002], and semi-supervised [Melville and Sindhwani, 2008]. Compared to supervised and unsupervised methods, semi-supervised methods for sentiment classification become more and more popular due to their making use of both the labeled and unlabeled data. This paper mainly focuses on semi-supervised methods for sentiment classification.

One kind of semi-supervised methods for sentiment classification is to utilize prior lexical knowledge in conjunction with the labeled and unlabeled data. For example, Document-Word Co-regularization for Semi-Supervised Sentiment Analysis analyzed the sentiment of documents and words based on a bipartite graph representation of the labeled and unlabeled data while Li et al. [2009] employed some simple update rules to make use of tri-factorization of the term-document matrix. It is rather common that such methods require a high-quality lexicon with the polarity of words properly defined [Melville and Sindhwani, 2008].

Another kind of semi-supervised methods for sentiment classification is to employ some bootstrap techniques, such as self-training [Yarowsky, 1995] and co-training [Blum and Mitchell, 1998]. Among them, co-training has been proven more effective than self-training [Wan, 2009; Huang et.al, 2010]. The key issue of applying co-training is to find a suitable set of different views. For instance, Co-Training for Cross-Lingual Sentiment Classification regarded two different languages (i.e., English and Chinese) as two views [Wan, 2009] while Employing Personal/Impersonal Views in Supervised and Semi-supervised Sentiment Classification considered personal and impersonal texts as two views [Huang et.al, 2010]. This paper employs the co-training technique and generates different views from random feature subspaces. Among others, Mine the Easy and Classify the Hard: Experiments with Automatic Sentiment Classification integrated various methods, such as spectral clustering, active learning, transductive learning, and ensemble learning, in semi-supervised sentiment classification [Dasgupta and Ng, 2009]. To our best knowledge, no existing semi-supervised methods consider the class imbalance problem in sentiment classification.

Semi-Supervised Learning for Semantic Relation Classification using Stratified Sampling Strategy explores several key issues in semi-supervised learning based on bootstrapping for semantic relation classification. The application of stratified sampling originated from statistics theory to the selection of the initial seed set contributes most to the performance improvement in the bootstrapping procedure. In addition, the more strata the training data is divided into, the better performance will be achieved. However, the augmentation of the labeled data using the stratified strategy fails to function effectively largely due to the unbalanced distribution of the confidently classified instances, rather than the stratified sampling strategy itself [Kong et.al, 2009].

Semi-supervised Learning for Relation Extraction integrate the advantages of SVM bootstrapping in learning critical instances and label propagation in capturing the manifold structure in both the labeled and unlabeled data, by first bootstrapping a moderate number of weighted support vectors through a co-training procedure from all the available data, and then applying label propagation algorithm via the bootstrapped support vectors [Li et.al, 2008].

Open Problems in Efficient Semi-supervised PAC Learning address semi-supervised learning for imbalanced sentiment classification. It adopts under-sampling to generate multiple sets of balanced initial training data and then propose a novel semi-supervised learning method based on random subspace generation which dynamically generates various subspaces in the iteration process to guarantee enough variation among the involved classifiers. Evaluation shows that semi-supervised method can successfully make use of the unlabeled data and that dynamic subspace generation significantly outperforms traditional static subspace generation [Balcan and Blum, n.d.].

2.3 Synthesis of the Study

Word Filtering is one of the most commonly used techniques in the recognition of the inappropriate expressions. Most of this is implemented via a word list and some regular expressions. It is not effective in grasping the context of inappropriateness in the expression causing more expressions to be falsely recognized.

The solutions using Lexical Syntactic features and Grammatical Relations play a big role in the recognition of determining inappropriate expressions. The said features help in identifying the usage in the sentence level, and then how it affects the person related in the expressions, with but lacks independence due to the loss of machine learning techniques. Bootstrapping approach for learning has been proven to be effective in learning, and modeling of linguistic data, though it is preferred to have forms of supervision rather than being unsupervised. There is a need for learning for the features of inappropriate expressions to further model the inappropriate language, in which it contains inappropriate expressions, thus coming up with a bootstrapping methodology.

CHAPTER III

RESEARCH METHODOLOGY

The aim of this chapter is to discuss the research design and methodology that was utilized in this study and research activities were be undertaken by the researchers. In order to describe the variety of research method and activities, research method, research paradigm, research design system architecture, data gathering procedure, and instrumentation will be systematically discussed.

3.1 Research Method Used

The experimental research method involves manipulating one variable to determine if changes in one variable causes changes in another variable. This method relies on controlled methods, random assignment and the manipulation of variables to test a hypothesis.

The system would not require respondents since the system used Experimental Research. Instead, the researchers would assess the performance of the system given implementation of algorithms and techniques. The system was tested by inputting a document containing inappropriate expressions and calculates its performance in terms of scores in accuracy.

3.2 Research Paradigm

The researchers would implement positivist way of approach. This is because Inappropriate Expressions Recognition requires lot of testing and observational analysis to ensure accurate and better result compare to the other studies. Series of testing and analysis of the system will be done to get the result with the highest level of accuracy.

In this way, the researchers had series of data that acted as an empirical evidence and later, was used to analyze the system accuracy and performance level.

3.3 System Architecture

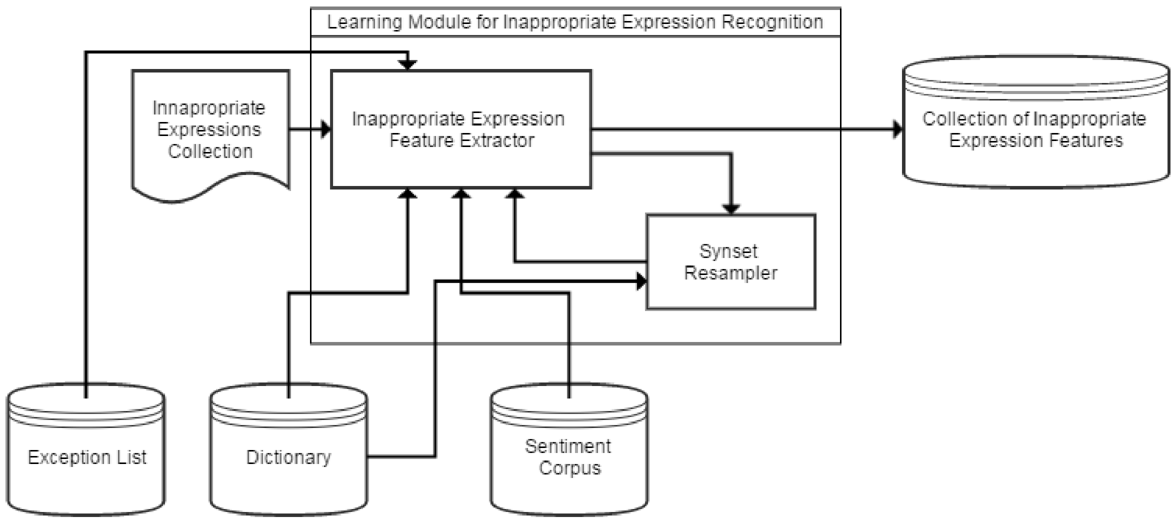


Figure 4 – Learning Module for Inappropriate Expressions Recognition

In the Training phase there is a Learning Module for inappropriate expressions. The learning module input consists of a file that contains lists of words that is deemed as Inappropriate. The learning module evaluates the Inappropriateness with basis of its polarity value in the Sentiment Corpus (which will be implemented via SentiWordnet) and the polarity of the definition of the word, in which it will be extracted in two dictionaries (via WordNet dictionary and Urban Dictionary website via Web Scraping) with the implementation of Naïve Bayes model. The Inappropriate expression back propagation will be done by extracting the feature in the definition that made the input inappropriate, and will be collected to the inappropriate expression features knowledge base. The synset resampling gets the synsets of the word and will undergo to the phases undergone by the original word. There is an exception list implemented to compensate and filter the noisy data descriptions of Urban Dictionary that causes false positives. The training module repeats this per word in the collection until all are evaluated and there are no more synsets to be resampled. After the Learning phase, the Learner Module offsets a threshold between on the mean and the global minima of the feature set as a computational borderline for inappropriate expressions. This learning is for the Unigram Expression trainin

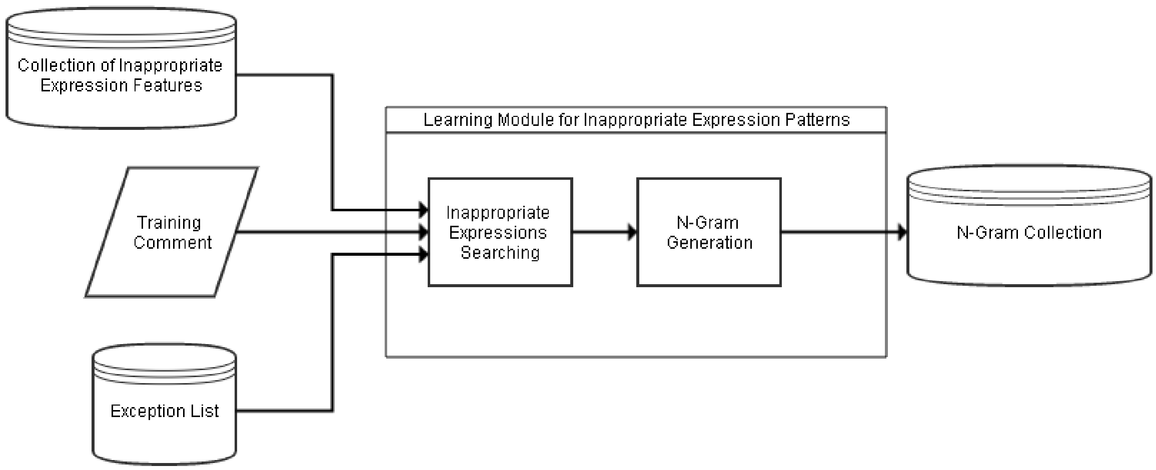


Figure 5 – Learning Module for Inappropriate Expressions Patterns

There is another Learning Module for inappropriate expressions which is used in relation to the neighboring words. This learning module input consists of a training comment and inappropriate expressions feature sets. The learning module finds inappropriate expressions on the input and tags them. There is an exception list implemented to remove the appropriate expressions that are tagged as inappropriate to filter the noisy data descriptions of Urban Dictionary that causes false positives. After tagging the inappropriate expressions, an N-Gram generator generates N-Grams from 2-Grams to 5-Grams based on the POS Tags of the neighboring grams.

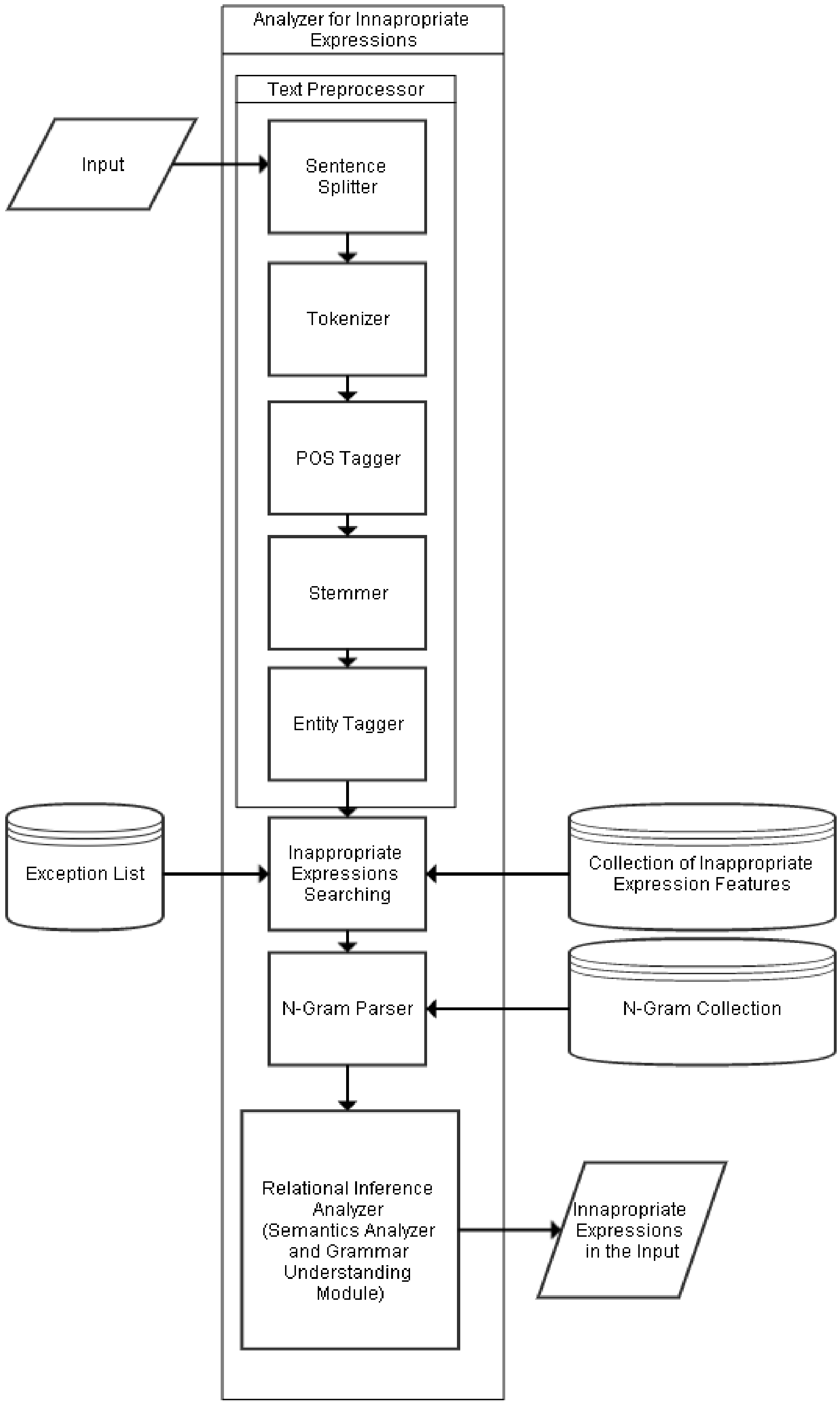


Figure 6 – Analyzer for Inappropriate Expressions

In the Analyzer Module, there will be an input of a comment. A document will undergo preprocessing. The preprocessing phase consists of Sentence splitting, Tokenization, Part-of-Speech Tagging, and Stemming (for the extraction of base form), and Entity Recognition. After undergoing preprocessing, for each sentence there will be a search for candidates in inappropriate expressions, which will be based on the collected features in the knowledge base. There is an exception list implemented to remove the appropriate expressions that are tagged as inappropriate to filter the noisy data descriptions of Urban Dictionary that causes false positives. Then the sentence will undergo to the N-Gram parsing to determine the probable usage of the inappropriate expressions in an inappropriate sense. After the parsing, the Relational Inference Analyzer determines the inappropriateness of the candidate words based each words' Lexical Syntactic Features and its grammar relations to the other existing words in the same input.

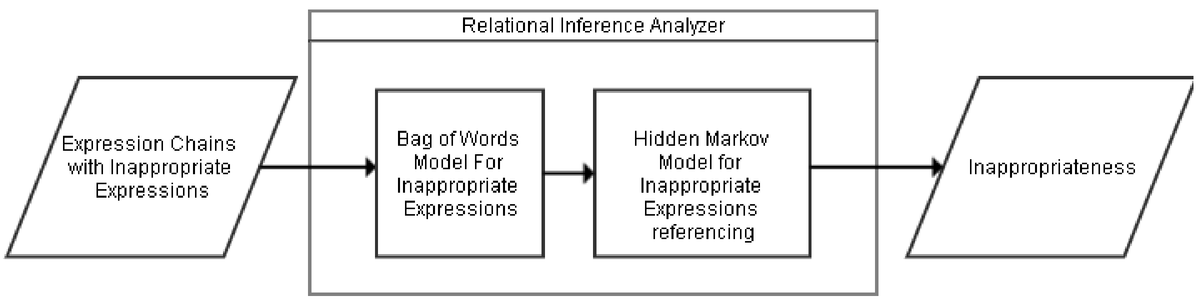


Figure 7 – Relational Inference Analyzer

The Relational Inference Analyzer is composed of multiple models that determines the Inappropriateness of the expression. First is the Bag of Words model that counts the candidate inappropriate expressions if they are at least 40% of the expression chain. The Hidden Markov Model determines candidate Inappropriate Expressions in relation to other inappropriate expressions and classified entities to determine the inappropriateness.

3.4 Sampling Technique

The researchers implemented purposive-quota sampling. Purposive-Quota Sampling is a non-representative subset of some larger population, and is constructed to serve a very specific need or purpose. The researchers deliberately set the proportions of levels or strata within the sample. This was generally done to insure the inclusion of a particular segment of the population.

The subjects were inappropriate expressions from 9gag comments and youtube.com comments and were selected just because they were the easiest to recruit for the study and the researcher did not consider selecting subjects that are representative of the entire population.

3.5 Sample Size of the Study

Sample size is an important concept in statistics and refers to the number of individual pieces of data collected. A statistic's sample size is important in determining the accuracy and reliability of the system.

In contrast to other researches which is most likely used people as their population, this study focused on objects as its focus. These objects that were obtained from 9gag and YouTube comments were set by the statistician to 500 due to unknown total population.

3.6 Description of Subjects

The subjects which are used in testing the model were 9gag and YouTube comments. The reason why the researchers chose those websites is because some of the users engage in that media was using English Language as a medium of communication. Another reason was because of publicity of the content of these sites. The researchers also observed that most of the users there were using inappropriate expressions in expressing their feelings, reactions and opinions.

3.7 Instrumentation

Instrumentation refers to the tools or means by which researchers attempt to measure variables or items of interest in the data collection process. The system was deployed as a Java Application with some Python scripts embedded in the system and the tools used in developing the system are WordNet, SentiWordNet, Stanford CoreNLP, MIT JWI Stemmer, requests module, and beautifulSoup. This tools served as the dictionary and was used in the pre-processing phase of the system. The system that would be dependent on the Latest Java Virtual Machine and Runtime Environment, and Python Interpreter.

The study utilized experiments to test its effectiveness on recognition of inappropriate expressions. So as the researchers had used experiment paper to identify the results of the tests conducted.

3.8 Data Gathering Procedure

The data gathered by the researchers came from the results in the experiments performed. The experimentation of the model was done by testing the performance in the recognition of Inappropriate Expressions and Appropriate Expressions.

There are two experiments performed in each file:

1) Testing for the Accuracy – This experiment was done by providing documents with inappropriate expressions in the model and testing it if it recognized them.

2) Testing for the Specificity - This experiment was done by providing documents without any inappropriate expressions or inappropriate expressions without inappropriate sense in the model and testing it if it avoided them.

3.9 Statistical Treatment

Statistical treatment consists of formulas that were used to answer what the problem states. The formulas used are as follows:

1. Specificity

   The performance of the recognition will be measured through the use of Specificity. Specificity is the rate of the results without the condition, which has a negative test result. NLP studies uses specificity to eliminate biased results in the system.

$$Specificity = \frac{TN}{TN + FP}$$

Where:

TN (True Negative) – the system and the expert correctly indicated that the input is appropriate

FP (False Positive) – System determined the input is Inappropriate, while the expert is appropriate

2. Harmonic Mean or F-Measure

The performance of the Recognition will be measured through the use of the Harmonic Mean, or f-measure. The f-measure is the weighted average of the values of the Precision and Recall. By multiplying the values by 2 and dividing it by the sum of the Precision and Recall, we can get the harmonic mean of the system. A high F1 score will imply a good performance of the system. The formula for the Harmonic mean is as follows:

$$Precision = \frac{TP}{TP + FP} \, x \, 100$$

$$Recall = \frac{TP}{TP + FN} \, x \, 100$$

$$F - Measure = \frac{2PR}{(P + R)}$$

Where:

TP (True Positive) – system and expert correctly determined the input is Inappropriate

FP (False Positive) – system determined the input is Inappropriate, the expert says otherwise

FN (False Negative) – system determined that the input is appropriate, the expert says otherwise

P = Precision – Percentage of identified expressions that are inappropriate.

R = Recall – Percentage of inappropriate expressions correctly identified.

CHAPTER IV

PRESENTATION, ANALYSIS AND INTERPRETATION OF DATA

This chapter presents how the proponents interpreted the results of the testing to come up with solutions to the problems raised in the study. It also shows the performance of the developed system Inappropriate Expressions Recognition using Bootstrapping as Semi-Supervised Learning in terms of the accuracy of the model and specificity of the model. The proponents underwent systematic testing, evaluation and analysis to assess the effectiveness of the developed model.

The following are the answers to the problem statement which were tabulated, interpreted and analyzed.

1. What is the performance analysis of the model in terms of:

    1.1 Recognition of Inappropriate expressions. (Accuracy of the Model)

    1.2 Recognition of Appropriate expressions. (Specificity of the Model)

The proponents gathered 500 YouTube and 9gag comments through the use of purposive-quota sampling due to its unknown total population. Gathered comments were divided into 50 .txt files which consists of 10 comments each. The files were tested individually and the results were tallied for each tagged inappropriate expression, such as number of inappropriate expressions that must be tagged by the system, number of inappropriate expressions that were correctly tagged by the system based on the expert (TP), number of inappropriate expressions that were correctly tagged by the system but not the expert (FP), number of inappropriate expressions that were correctly tagged by expert but not the system (FN) and number of appropriate expressions that were correctly tagged by the expert and system (TN). From these classifications of results, the researchers computed the Precision by dividing the number of inappropriate expressions that were correctly tagged by the system and expert over the sum of number of inappropriate expressions that were correctly tagged by the system and expert and the number of inappropriate expressions that were correctly tagged by the system but not the expert (TP/(TP+FP)), the Recall by dividing the number of inappropriate expressions that were correctly tagged by the system and the expert over the sum of the number of inappropriate expressions that were correctly tagged by the system and the expert and number of inappropriate expressions that were correctly tagged by the expert but not the system (TP/(TP+FN)), the F-Measure by dividing the product of precision and recall to two over the sum of precision and recall (2PR/(P+R)) and the Specificity which is computed by number of appropriate expressions that were correctly tagged by the expert and system over the sum of number of appropriate expressions that were correctly tagged by the expert and system and number of inappropriate expressions that were correctly tagged by the system but not the expert (TN/(TN+FP)). Tables 1-4 show the assessment.

4.1 Results for Precision

Table 1 - Overall Evaluation in Terms of Precision

CLASSIFICATION	TOTAL NUMBER	PRECISION (TP/(TP+FP))
TP	389	**73.12**
FP	143	

Table 1 presents the overall performance of the system in terms of Precision. Precision is the percentage of identified expressions that are inappropriate. The overall precision of the 50 files tested was 73.12% because according to the researchers, some of the expressions that are contextually appropriate are still recognized as inappropriate in context. The primary reason for such was seen in Figure 8, it is because some of the Lexical Syntactic features of the Expression are noise data from definitions of UrbanDictionary.

Figure 8 – Sample comment with False Positive Result

4.2 Results for Recall

Table 2 - Overall Evaluation in Terms of Recall

CLASSIFICATION	TOTAL NUMBER	RECALL (TP/(TP+FN))
TP	389	**66.84**
FN	193	

Table 2, presents the overall performance of the system in terms of Recall. Recall is the percentage of inappropriate expressions that are correctly identified. To compute this, the researchers were accompanied by an English Teacher as their expert in evaluating each comment found in each file. The overall recall of the 50 files tested was 66.84% because according to the researchers, some of the expressions that are contextually inappropriate are still unrecognized by the system. The primary reason for such was seen in Figure 9, it is because grammar relations are not fully established due to the uncontrolled scoping of the phrase-level orientation from the N-Gram Language Model.

Figure 9 – Sample Result with False Negative Result

4.3 Results for F-Measure

Table 3 - Overall Evaluation in Terms of F-Measure

CLASSIFICATION				PRECISION	RECALL	F-MEASURE (2PR/(P+R))
TP	FP	TN	FN			
389	143	5662	193	73.12	66.84	**69.84**

Table 3, shows the overall performance of Inappropriate Expressions Recognition using Bootstrapping as Semi-Supervised Learning in terms of F-Measure. The average F-

Measure of the 50 files tested was 69.84% because based on the previous tables, the recognizer produced low percentage of recall which affected the performance of the system.

4.4 Results for Specificity

Table 4 - Overall in Terms of Specificity

CLASSIFICATION	TOTAL NUMBER	SPECIFICITY (TN/(TN+FP))
TN	5662	**97.50**
FP	143	

Table 4, shows the overall performance of Inappropriate Expressions Recognition using Bootstrapping as Semi-Supervised Learning in terms of Specificity. Specificity is the rate of the results without the condition, which has a negative test result. The average Specificity of the 50 files tested was 97.50% because contextually appropriate expressions that were not tagged by both system and expert were greater than the tagged inappropriate expressions that are not contextually inappropriate to the expert.

CHAPTER V

SUMMARY OF FINDINGS, CONCLUSION AND RECOMMENDATION

This study was conducted for the purpose of determining the performance of the system Inappropriate Expressions Recognition using Bootstrapping as Semi-Supervised Learning. Since the research was experimental, the system went through several testing and evaluation to test the accuracy and ambiguity of the generated output with numerous test data. This chapter presents the summary of findings, conclusion and the recommendations.

5.1 Summary of Findings:

This study aimed the creation of inappropriate expressions recognizer that recognizes English comments found in 9gag and YouTube and tags inappropriate expressions based on its context. For this, the researchers used the experimental research method. The system went through several testing and evaluation to test the accuracy and ambiguity of the generated output in terms of Precision, Recall, F-Measure and Specificity. The statistical tool used by the researchers to test the accuracy of the developed system was the F-Measure and to eliminate bias, the researchers used Specificity.

The following were the findings of the study based on the systematic testing, evaluation and analysis:

Based on the statement of the problem mentioned in Chapter 1 "What is the performance analysis of the model in terms of: 1.1 Recognition of Inappropriate expressions (Accuracy of the Model) and 1.2 Recognition of Appropriate expressions. (Specificity of the Model)" the following results were computed:

CLASSIFICATION				**PRECISION**	**RECALL**	**F-MEASURE (2PR/P+R)**
TP	**FP**	**TN**	**FN**			
389	143	5662	193	**73.12**	**66.84**	**69.84**

CLASSIFICATION	TOTAL NUMBER	SPECIFICITY (TN/TN+FP)
TN	5662	**97.50**
FP	143	

The summary assessment of the performance of the system based on 50 files tested in terms of Precision, Recall, F-Measure and Specificity were computed as 73.12%, 66.84%, 69.84% and 97.50% respectively.

5.2 Conclusion:

Based from the findings of the study entitled "Inappropriate Expressions using Bootstrapping as Semi-Supervised Learning" the proponents reached the following conclusions through the series testing and evaluation:

1. The overall performance of the developed system in terms of F-Measure is 70% and a Specificity of 98%.
2. The Exception List has provided a workaround against the noisy definitions of UrbanDictionary that causes false positives.
3. Bootstrapping has been an effective Semi-Supervised learning schema for Inappropriate Expressions.
4. The performance of the system can be improved if more feature functions were fed into the system.

5.3 Recommendation:

The following suggestions might be helpful for those future researchers who will also specialize in any topic relating to Inappropriate Expressions Recognition:

1. Compare the significant difference of the developed Inappropriate Expressions Recognizer that utilizes different algorithm to further emphasize the usefulness of the Bootstrapping approach.
2. This system can be improved by broadening the feature word neighboring to sentence neighboring in order to determine correctly the context of the detected inappropriate expressions.

3. There is a need for a supervised N-Gram Modeling to model the inappropriate expressions in a proper phrase level scoping.
4. There should be a control of which UrbanDictionary definitions to be used in avoiding noise of false positives and false negatives.
5. There is a need for a methodology to handle the noisy definitions of UrbanDictionary to remove the need of an exception list.
6. There may be a need of usage of other dictionaries like Merriam-Webster's Dictionary and Oxford Dictionary (Merriam-Webster's Dictionary and Oxford's dictionary needs legal permissions before usage).
7. There may be a need for the recognition of multiple word idiomatic inappropriate expressions.

BIBLIOGRAPHY

Print Publications

Balcan, M. and Blum, M. Open Problems in Efficient Semi-Supervised. *In Proceedings of National Science Foundation Grant CCF-0514922 and a Google Research Grant* (n.d.).

Balahur, M., Lloret, E., Montoyo, A. and Palomar M. Towards Building a Competitive Opinion Summarization System. *In Proceedings of NAACL-09 Student Research Workshop and Doctoral Consortium* (2009), pp. 72-77.

Baum, L. E. and Petrie, T. Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *The Annals of Mathematical Statistics*, 37, 6 (November 1966), pp. 1554-1563.

Bellmore, A., Xu, J.M. and Zhu, X. Fast Learning for Sentiment Analysis on Bullying. *In Proceedings of ACM 978-1-4503-1543-2/12/08* (August 2012).

Blum, A. and Mitchell, T. Combining Labeled and Unlabeled Data with Co-Training.*In Proceedings of the Eleventh Annual Conference on Computational Learning Theory* (1998), pp. 92-100.

Bowyer, K., Chawla, N., Hall, L. and Kegelmeyer, W. SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*, 16 (2002), pp. 321-357.

Boudaren et. al., Unsupervised segmentation of random discrete data hidden with switching noise distributions. *IEEE Signal Processing Letters*, 19, 10 (October 2012), pp. 619-622.

Cardie, C., Wiebe, J. and Wilson T. Anotating Expressions of Opinions and Emotions in Language. *Language Resources and Evaluation*, 39, 2/3 (2005).

Chen, Y., Xu, H., Zhou, Y. and Sencun, Z. Detecting Offensive Language in Social Media to Protect Adolescent Online Safety. *In Proceedings of PennState College of Information Sciences and Technology* (2012).

Cheng, J., Hu, M. and Liu, B. Opinion Observer: Analyzing and Comparing Opinions on the Web. *In Proceedings of WWW-05* (2005), pp.342-351.

Cornog, M. Naming Sexual Body Parts: Preliminary Patterns and Implications. *The Journal of Sex Research*, 22, 3 (August 1986), pp. 393-398.

Cui, H., Datar, M. and Mittal, V. Comparative Experiments in Sentiment Classification for Online Product Reviews. *In Proceedings of AAAI-06* (2006), pp. 1265-1270.

Dasgupta, S. and Ng, V. Mine the Easy and Classify the Hard: Experiments with Automatic Sentiment Classification. *In Proceedings of ACL-IJCNLP-09* (2009), pp. 701-709.

Duin R. and Juszczak, P. Uncertainty Sampling Methods for One-Class Classifiers. *In Proceedings of ICML-03, Workshop on Learning with Imbalanced Data Sets II* (2003), pp. 81-88.

Fan, B., Hong, J., Rose, C., Wang, L. and Xiang, G. Detecting Offensive Tweets via Topical Feature Discovery over a Large Scale Twitter Corpus. *In Proceedings of ACM 978-1-4503-1156-4/12/10* (2012).

Finkel, J.R., Grenager, T. and Manning, C. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. *43rd Annual Meeting of the Association for Computational Linguistics* (2005), pp. 363–370.

Huang, C., Lee, S., Li, S. and Zhou, G. Employing Personal/Impersonal Views in Supervised and Semi-supervised Sentiment Classification. *In Proceedings of ACL-10* (2010), pp. 414-423.

Kong, F., Qian, L., Xhou, G. and Zhu, Q. Semi-Supervised Learning for Semantic Relation Classification using. *In Proceedings of 2009 Conference on Empirical Methods in Natural Language Processing*, 6, 7 (August 2009), pp. 1437-1445.

Lee, L., Pang, B. and Vaithyanathan S. Thumbs Up? Sentiment Classification using Machine Learning Techniques. *In Proceedings of EMNLP-02* (2002), pp. 79-86.

Li, J.H., Qian, L.H., Zhou, G.D. and Zhu, Q.M. Semi-supervised Learning for Relation Extration. *In Proceedings of the 3rd International Joint Conference on Natural Language Processing (IJCNLP-2008)*, 7, 12 (January 2008), pages 32-38.

Liang, P. Semi-Supervised Learning for Natural Language. *In Proceedings of Massachusetts Institute of Technology* (May 2005), pp. 13-73.

Lin, Dekang and Wu, Xiaoyun. Phrase Clustering for discriminative learning. *Annual Meeting of the ACL and IJCNLP* (2009), pp. 1030-1038.

Liu, X. and Zhou, Z. Training Cost-Sensitive Neural Networks with Methods Addressing the Class Imbalance Problem. *IEEE Transaction on Knowledge and Data Engineering*, 18 (2006), pp.63-77.

Melville, P. and Sindhwani, V. Document-Word Co-regularization for Semi Supervised Sentiment Analysis. *In Proceedings of ICDM-08* (2008), pp. 1025 1030.

Norvig, P. and Russell, S. Artificial Intelligence: A Modern Approach (2nd ed.). *Prentice Hall* (2003). ISBN 978-0137903955.

Rish, I. An Empirical Study of the Naive Bayes Classifier (PDF). *IJCAI Workshop on Empirical Methods in AI* (2001).

Satish, L. Use of Hidden Markov for Partial Discharge Pattern Classification. *IEEE Transactions on Dielectrics and Electrical Insulation* (April 2003).

Stratonovich, R.L. Condiitonal Markov Processes. *Theory of Probability and its Application*, 5, 2 (1960), pp. 156-178.

Turney, P. Thumbs up or Thumbs down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *In Proceedings of ACL-02*, (2002), pp. 417-424.

Wan, X. Co-taining for Cross-Lingual Sentiment Classification. *In Proceedings ofACL IJCNLP-09* (2009), pp. 235-243.

Sivic, J. Efficient Visual Search of Videos Cast as Textual Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31, 4 (April 2009), pp. 591-605.

Wells, G. Talk about Text: Where Literacy is Learned and Taught. *Curriculum Inquiry*, 20, 4(1990), pp. 369-405.

Xu, Z. and Sencun Z. Filtering Offensive Language in Online Communities using Grammatical Relations. *In Proceedings of CEAS 2010 - Seventh Annual Collaboration, Electronic Messaging, Anti-abuse and Spam Conference* (July 2010).

Yarowsky, D. Unsupervised Word Sense Disambiguation Rivaling Supervised Methods. *In Proceedings of ACL-05* (1995), pp. 189-196.

Other Printed Sources

Brants, T. and Franz, A. 2006. All our N-gram are belong to you. Retrieved December 16, 2011

Dunning, T. *Statistical Identification of Language*. New York State University, 1994.

Efron, B. and Tibshirani, R.J. *Chapter 8: The Bootstrap*. 1993.

Goodenough, F.L. Anger in Young Children. *Wesport, CT: Greenwood Press* (1931).

Harris, Z. Distributional Structure. *Word 10*, 2, 3 (1954).

Hong, L., Xue, Z. and Yin, D. Detection of Harassment on Web 2.0. *In Proceedings of Lehigh University: Computer Science and Engineering* (April 2009).

Jay, T. *Why We Curse?* John Benjamins Publishing Company, North Adams, Massachusetts, 2009.

Northman, J., et.al. Learning multilingual named entity recognition. *Artificial Intelligence 194* (2013), pp. 151-175.

Electronic Sources

Chapter 4: The Original Bootstrap Method. n.d, from VirginiaTech: http://scholar.lib.vt.edu/theses/available/etd-6169714555/unrestricted/Ch4.pdf

"Definition of Profanity", Retrieved from on August 08, 2014 from Merriam-Webster Online Dictionary: http://www.merriam-webster.com/dictionary/profanity

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

FileSharingTalk. Retrieved October 1, 2006 from *When the **** did we get a Word Filter?*: http://filesharingtalk.com/threads/88125-When-the-****-did-we-get-a-wordfilter

Grohol, J. 2009. PsychCentral. Retrieved March 30, 2009 from *Why Do We Swear?*: http://psychcentral.com/blog/archives/2009/03/30/why-do-we-swear/

Inappropriate Content. Retrieved 18 September, 2015 from Ed422 Cybersafety 101: https://sites.google.com/site/nbushra

Semi-supervised Learning. Retrieved September 18, 2015 from *Wikipedia: The Free Encyclopedia*: https://en.wikipedia.org/wiki/Semi-supervised_learning

Sheerin, J. 2010. BBC News. Retrieved March 29, 2010 from *How spam filters dictated Canadian magazine's fate*: http://news.bbc.co.uk/2/hi/technology/8528672.stm

Watson. from *IBM Research*: http://www.research.ibm.com/cognitivecomputing/watson/ index.shtml#fbid=_V9cRK7OhZC

Where Did IBM's Super Computer Watson Learn to Swear?. From *How-to-Geek:* *http://www.howtogeek.com/trivia/where-did-ibms-super-computer-watson-learn-to-swear/*

APPENDIX A

SOFTWARE EVALUATION TOOL

SOFTWARE EVALUATION TOOL

Polytechnic University of the Philippines
Sta. Mesa, Manila
College of Computer and Information Sciences

Inappropriate Expressions Recognition using Bootstrapping as Semi-Supervised Learning

Date: ___________________

File No.	No. of Tokens	TP	FP	TN	FN	Precision	Recall	F-Measure	Specificity

Legend:

- TP (True Positive) - system and expert correctly determined the input is Inappropriate
- FP (False Positive) – System determined the input is Inappropriate, while the expert is appropriate
- FN (False Negative) – system determined that the input is appropriate, the expert says otherwise
- TN (True Negative) – the system and the expert correctly indicated that the input is appropriate

APPENDIX B

PSEUDOCODE

PSEUDOCODE

Analyzer for Inappropriate Expressions

```
function inappExp(String comment):
    List<Expression> expressions;
    String[] sentences=sentenceSplit(comment);
    foreach(String sentence in sentences):
        String[] tokens=tokenize(sentence);
        foreach(String token in tokens):
            expressions.add(token);
        generatePOS(expressions);
        generateNER(expressions);
        foreach(Expression expression in expressions):
            extractInappnessFeatures(expression);
        List<Expression[]> parsedExpressions=NGramParse(expressions):
        foreach(parsedExpression in parsedExpressions):
            BagOfWordsInappropriateness(parsedExpression);
            HiddenMarkovInappropriateness(parsedExpression);
            if(Inappropriate(parsedExpression)):
                foreach(expression in expressions):
                    if(inappropriate(expression)):

    output("<Inapp>"+expression.word+"</Inapp> ");
                    else
                        output(expression.word+" ");
```

Word Learn

```
function inappExpWordLearn(String word):
    List<String>definitionList;
    AddWordNetDefinitions(word,definitionList);
    AddUrbanDictionaryDefinitions(word,definitionList);
    foreach(String definition in definitionList):
        String[] tokens=tokenize(definition);
        foreach(String token in tokens):
            learnInappnessFeature(token);
```

N-Gram Learn

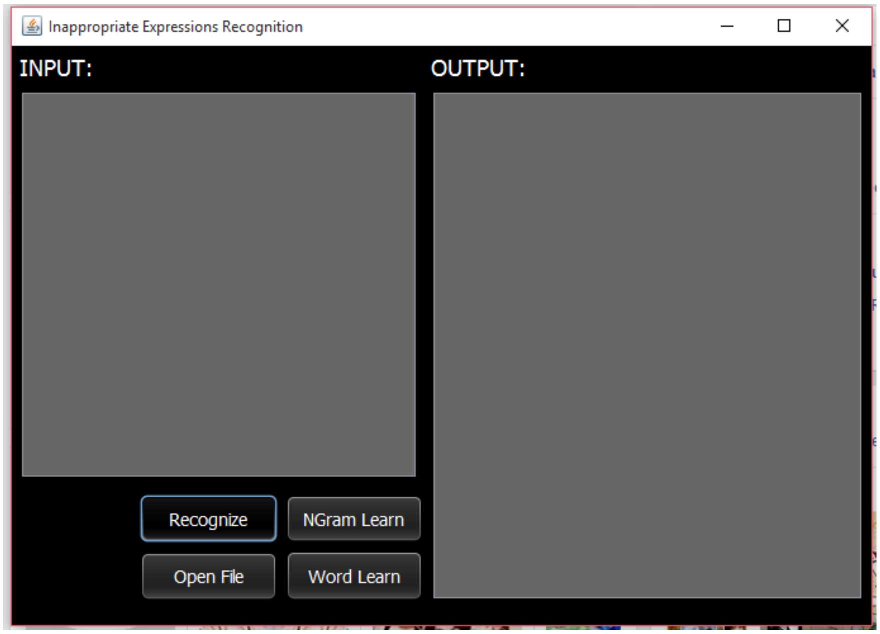
```
function inappExpNGramLearn(String comment):
    List<Expression> expressions;
    String[] sentences=sentenceSplit(comment);
```

```
foreach(String sentence in sentences):
	String[] tokens=tokenize(sentence);
	foreach(String token in tokens):
		expressions.add(token);
	generatePOS(expressions);
	generateNER(expressions);
	foreach(Expression expression in expressions):
		extractInappnessFeatures(expression);
	NGramLearn(expressions);
```

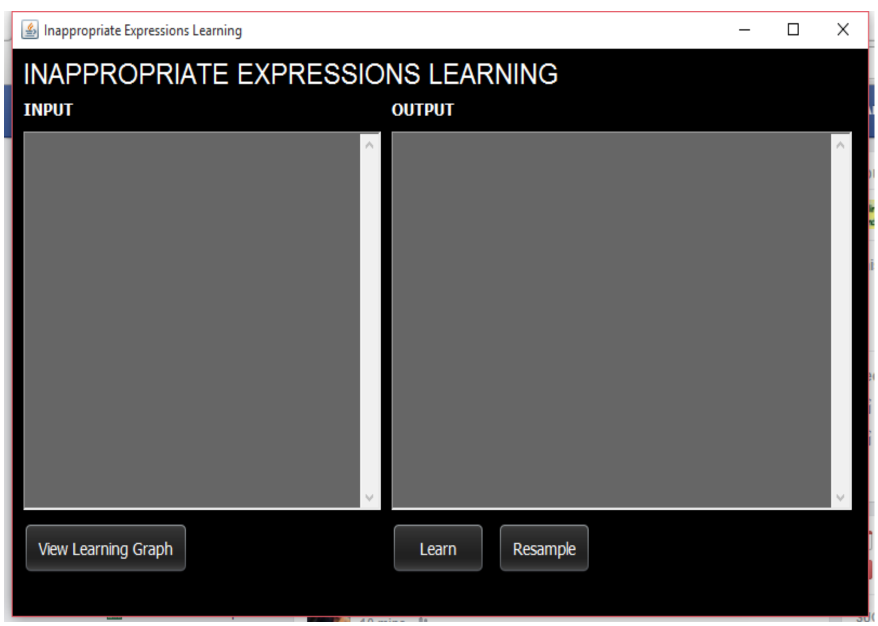
APPENDIX C

SCREENSHOTS

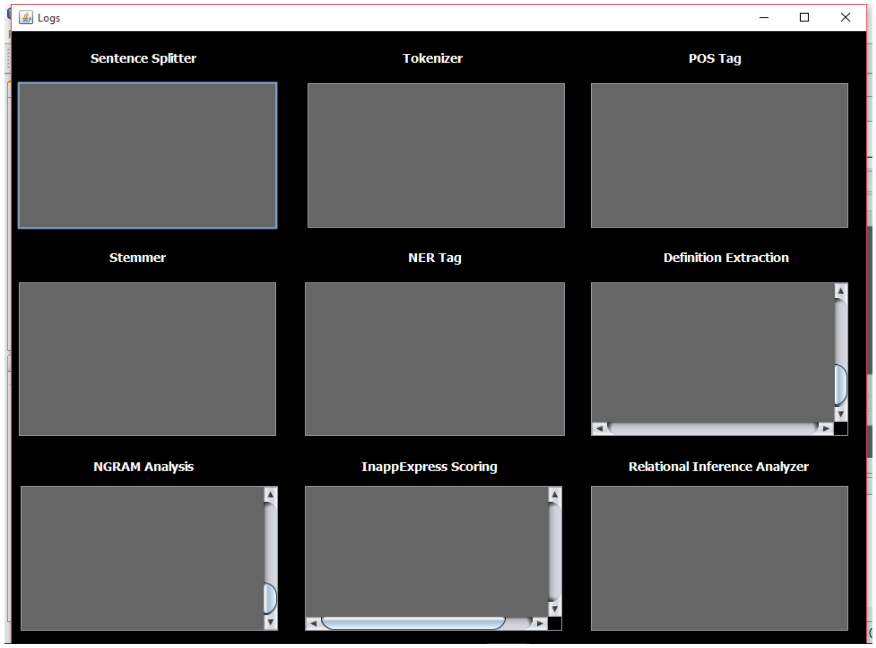
SCREENSHOTS



Program Environment



Learning Environment


Logs
Sentence Splitter
Tokenizer
POS Tag
Stemmer
NER Tag
Definition Extraction
NGRAM Analysis
InappExpress Scoring
Relational Inference Analyzer

Logs

APPENDIX D

LIST OF SAMPLE DATA

LIST OF SAMPLE DATA

File No. 1

The main reason I don't want children is because I don't want to go through pregnancy. Morning sickness, loss of bladder control, you have to pee constantly, weird cravings, your mental ability decreases, you gain weight, your figure goes to shit....If someone showed up at my door with a newborn and handed it to me, I would raise that fucker like a champ. But if one month my period is late, and it turns out I'm pregnant, I would freak the fuck out.

I cant imagine how women can deal with the pain when they pull this shit out of their juju , oh god . :O

They are high as fuck when they give birth. Helps a lot.

And they still won't let us fuck them in the ass because it hurts. Bullshit.

Scumbag baby

Where the fuck goes the all the intestines?

Female human got infected via sex and the parasite grow in their body until they hatch out.

Nice...boobs gets bigger

Boobs will definitely grow more than that

Fuck humanity, I would never wish that pain on anybody. Don't worry, future imaginary wifey, we can adopt

File No. 2

All i can look at are the boobs

I love the boob growing

I may sound like an asshole, but I would be kind of relieved if my future wife couldn't have kids. I would love to be a dad someday, but the thought of my wife going through something like that and possibly dying while giving birth really terrifies me

that some scary shit dude

Poor intestine, it is just like, "hey what's going on, uterus plz, uterus stahp, guys some space please, well fuck"

I'm a simple man, I see a drawing of a baby growing inside naked woman, I fap

He's fucking 21 and she's 30 something cmon?

Miranda is also older but shit happened and now she is not married anymore...?

Milf bro?

Yeah but she got kids and Justin can fuck anybody else lol?

File No. 3

she fucking fine. if he want to fuck her, that's all the reason he needs?

Well think about this way he fucks one of them why would he pick kourtney now he can't fuck any of the sisters. I would have gone with Kendall or khole on some real shit?

So that pussy is is still tight?

I'll take that back that pussy is lose?

her anus probably not tight anymore either?
Still fucking hot lol?
Lol you're a dumbass. Why would he even photoshop that. It's not like he was trying to make himself skinnier or some shit. Use common sense?
And his arms aren't through his sleeves in his jacket retards?
I bet you got a fat juicy butt?
well i like milfs because I can actually go balls deep, and the skin is softer rather than firm like a 20 yr old.?

File No. 4
are u fucking gay? Lol?
your a faggot?
Each year I hate Justin a little less. If he smashed her I'll never say nothing bad about him again.?
this is the stupidest shit ive seen today , i even went to school and there is a lot of stupid shit there.?
I agree - this was pathetic, desperate and sad...?
Every time she has sex she brings home the evidence. We will know in 9 months.?
Besides that wasnt he fucking kendel a while back? Lord knows they a bunch of dirty hoes?
That's so nasty she has 3 kids and she's 36 obviously she's the easiest to smash lol?
I'd fuck her any day . She a MILF?
how is Kourtney the easiest to smash ??

File No. 5
sometimes i wonder why i even check back here once a month, like a dont really care and it just enrages me that people care so much.?
So besides her leaving a hotel that she very well could have stayed at, where is the evidence that they are hooking up? Clearly that picture is now irrelevant... TMZ is literally such bullshit?
Then why tf did u fucking say "but it's true" you fucking fucks?
his people say he started hitting on her... this video is bullshit?
maybe he's tired of basic generic hot bitches and wants to go milf?
so they hooked up..okay. congrats to the Canadian kid who fucked a milf.?
She's denying it so she won't look like a hypocrite.?
who cares if they are fucking god damn TMZ you on the Kardashians dick?
he smashin?
We all know if Kim can still lie saying her ass is real and kylie about them lips, Khloe increasing her ass and saying its due to squats, this i can also believe, The End?

File No. 6
ikr she's still quite bangin for someone with 3 kids. damn ;)?

This is so stupid. Get a life of your own tmz.
I blame faggot media outlets like TMZ for keeping this no talent midget relevant for this long. Back in the day Vanilla Ice was gone in a couple of months, nowadays that douche would have a ten year run, he at least he had a hit record.?
If he can bone the tightest pussies in the world, why in the world would he want a worn out stretched pussy like KK's??
This is fucking ridiculous. then again, Kourtney was stupid enough to be with "lord" disick for so long. Why not.?
So because they were in the same hotel and a stupid Instagram and your "sources", they're hooking up. Outstanding logic TMZ?
why would she sleep with this pussy, what a downgrade from scott. This kid is a punk wannabe.?
Justin hooked up with Kendall and kourteny
Milf chaser,lol?
out of all people why would she date Justin bebber, if she did so must have been high?

File No. 7
Man I swear tmz is just like a high school rumour club?
It's photoshop. I mean I don't like Bieber, but he sure did fool you guys, hard.?
This is stupider than the love triangle happening at my high school right now.?
Fuck outta here with tht boolshit don't report this shit?
bitch hooking up bitch
Yummy Mummy!?
I hate jasmin this motherfucker BITCH
Ever since Biebers nudes leaked, she hopped on that shit.?
You're the same nigga who'd eat her cigarette filled shit if she said you could email her. lmfao?
WELL? did he ever get that DATE? UGHH I HATE CLIFFHANGERS?

File No. 8
Lana is fine as shit, idk why I always assumed she was like a 40+ year old lady?
if you wear socks with sandals, you deserve to be run the fuck over.?
number 2, deserved asshole.?
Like someone else said in the comments, Lana is fine, but for whatever reason one would assume she's 40. Not saying 40 year old women are not attractive, but clearly Lana is not 40, she just somehow gives of that vibe and I don't know why.?
I FUCKING LOVE LANA DEL REYEY?
this guy tryna blame her for driving there.. don't be a fucking creep and follow her everywhere then, hope it broke tbh?
he's weird. he's here. and he's queer. Lol jk?

That last one wasn't awkward, that was fucking badass. The camera guy is living the dream man.?
Lana's mine screw your Camera Guy?
lucky mother fucker?

File No. 9
Why are white guys so obsesesed with the size of a black mans dick and curious to know how he showers?...?
start banging out more videos like this, rather watch a 5min video then a 1min new celeb video, i want bloopers and funny shit lol?
DAMN BRO give me her email man SHES GORGEOUS!!!?
having a ass plus money. well money for an ass lol?
Kris didn't mind when he was giving her the big black dick when she was 17 why would she care now?
Holy fuck... It just hit me.. Kim is 30+ & was more my older brothers generation & this bitch (Kylie) is taking her place so when I get 23 or some shit. The world will be all up Kylie's ass like they were up Kim's
Kim got ass shots and lazer work. Shit looks like a diaper booty?
Her ass is real she just wears padded underwear sometimes?
just jealous you don't have an ass.?
Ikr she was fat.?

File No. 10
Fuck now im horny! Back to pornhub?
Amber rose is a hoe?
just like her big bald head?
you stupid if you think that , her tits saggy asf only natural tits can. Her ass has cellulite?
she was a stripper at 14,15 all kind of men have been sicking them and she had a child so they are going to a bit floppy. AND TO BE HONEST ALL I SEE IS PERFECTLY MELLONS?
Kims ass is fake like the rest of her body and life. Amber 100% natural. Anybody that says Kim, ya must like fake better smh?
Amber wins because hers is real!!! These Kardashian dicks riders knows its true
u fuckin nerd Amber and Kim are better skinny bitches are boring?
Where's the black ladies at??? I know a list of famous black women who beat most of these girls. Smh. So disappointed
I notice the only guys who are into huge butts are blacks and hispanics. White men usually prefer their women slim and petite.?

APPENDIX E

INDIVIDUAL RESULTS

INDIVIDUAL RESULTS

File No.	No. of Tokens	TP	FP	TN	FN	Precision	Recall	F-Measure	Specificity
1	129	7	1	115	6	87.5	53.85	66.67	99.14
2	146	7	0	135	4	100	63.64	77.78	100
3	146	10	0	131	5	100	66.67	80	100
4	164	10	3	139	12	76.92	45.45	57.14	97.89
5	179	15	0	161	3	100	83.33	90.91	100
6	206	5	2	192	7	71.43	41.67	52.63	98.97
7	126	12	1	112	1	92.31	90.31	90.31	99.12
8	176	5	2	166	3	71.43	62.50	66.67	98.81
9	143	10	4	128	1	71.43	90.91	80	96.97
10	188	5	9	162	2	35.71	71.43	47.62	94.74
11	228	11	5	212	2	68.75	84.62	75.86	97.70
12	94	6	3	80	5	66.67	54.55	60	96.39
13	323	11	1	305	6	91.67	64.71	75.86	99.67
14	152	11	1	138	2	91.67	84.62	88	99.28
15	74	12	1	60	1	92.31	92.31	92.31	98.36
16	120	11	0	107	2	100	84.62	91.67	100
17	192	7	0	178	7	100	50	66.67	100
18	103	11	1	89	2	91.67	84.62	88	98.89

19	57	6	0	46	5	100	54.55	70.59	100
20	62	15	0	44	3	100	83.33	90.91	100
21	69	11	0	57	1	100	91.67	95.65	100
22	84	8	1	71	4	88.89	66.67	76.19	98.61
23	58	16	0	41	2	100	88.89	94.12	100
24	111	9	3	98	1	75	90	81.82	97.03
25	134	11	1	118	4	91.67	73.33	81.48	99.16
26	209	6	9	191	4	40	60	48	95.5
27	126	1	6	117	2	14.29	33.33	20	95.12
28	105	7	3	94	1	70	87.50	77.78	96.91
29	122	6	4	108	4	60	60	60	96.43
30	129	3	7	117	2	30	60	40	94.35
31	104	2	5	92	4	28.57	33.33	30.77	94.85
32	115	8	5	102	0	61.54	100	76.19	95.33
33	109	3	10	91	5	23.08	37.5	28.57	90.10
34	160	4	8	144	4	33.33	50	40	94.74
35	112	8	6	97	1	57.14	88.89	88.89	94.17
36	100	10	10	80	0	50	100	66.67	88.89
37	115	3	5	103	4	37.5	42.86	40	95.37
38	78	3	3	66	6	50	33.33	40	95.65
39	127	6	9	111	1	40	85.71	54.54	92.5

40	132	2	7	122	1	22.22	66.67	33.33	94.57
41	138	8	3	123	4	72.72	66.67	69.57	97.62
42	117	6	1	101	9	85.71	40	54.55	91.82
43	127	8	2	114	3	80	72.73	76.19	99.02
44	117	7	1	97	12	87.5	36.84	51.85	98.98
45	76	6	0	64	6	100	50	66.66	100
46	73	8	0	58	7	100	53.33	69.67	100
47	81	10	0	68	3	100	76.92	86.96	100
48	85	3	0	74	8	100	27.27	42.86	100
49	83	8	0	72	3	100	72.73	84.21	100
50	90	8	0	74	8	100	50	66.67	100

APPENDIX F

IMPLEMENTATION REPORT

IMPLEMENTATION REPORT

INTRODUCTION

Inappropriate Expressions is one of the problems in a behavioral sense. Inappropriate Expressions mostly causes problems in literary management like cyber bullying, and exposure of children to other textual data that may cause other interests like crime, sex, etc.

The problem is, due to the inherent ambiguity of the language, there is a hard time to recognize the real meaning if whether the expression gets inappropriate, making it harder to be recognized. The solution to be used in this study is a machine learning methodology, which is bootstrapping, which is designed to model the Inappropriate language, in which embodies Inappropriate Expressions. With this solution, users may find inappropriate expressions in textual data, which can be used as a tool for prevention of the exposure of the inappropriate expressions to those who are not concerned.

Inappropriate Expressions Recognition using Bootstrapping as Semi-Supervised Learning is a tool that tags inappropriate expressions present in each comments in .txt files. Words that does not have tags are considered as appropriate expressions, which means that, it is not inappropriate in context. The model accepts comments in pure English Language. YouTube and 9gag comments are the subjects in this study. This also models the sentence-level context analysis to identify the inappropriateness of an expression with the use of Lexical Syntactic Features and Grammar Relations as a support to the said computational model that is used to solve the problem of modeling the inappropriate language.

PROBLEM STATEMENT

The following problems are to be solved after the implementation:

1. What is the performance analysis of the model in terms of:

    1.1 Recognition of Inappropriate expressions. (Accuracy of the Model)

    1.2 Recognition of Appropriate expressions. (Specificity of the Model)

RESPONDENTS/SUBJECTS

The subjects which are used in testing the model are 9gag and YouTube comments. The reason why the researchers chose those websites is because some of the users engage in that media is using English Language as a medium of communication. Another reason is because of publicity of the content of these sites.

These objects that were obtained from 9gag and YouTube comments were set by statistician to 500 due to unknown total population.

TIME FRAME

3rd - 4th week of December	1st - 2nd week of January	3rd – 4th week of January	1st week of February	2nd week of February
Data Gathering	Data Gathering			
		System Tuning		
			Final System Testing	
				Finalization of Results

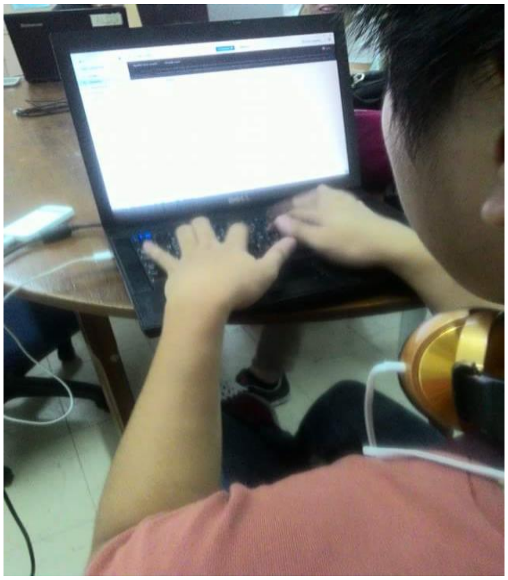
IMPLEMENTATION PROCEDURES

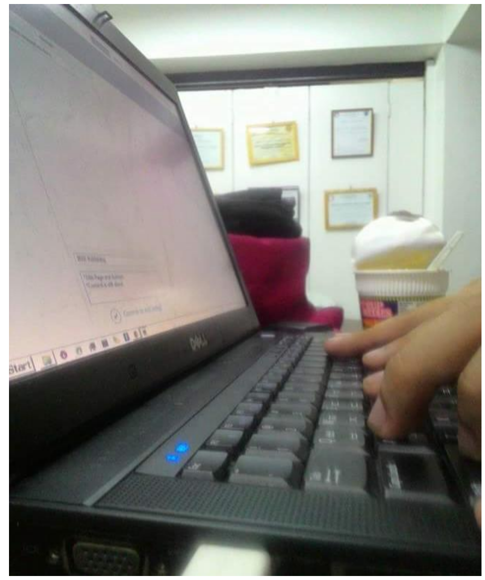
The proponents gathered 500 YouTube and 9gag comments through the use of purposive-quota sampling due to its unknown total population. Gathered comments were divided into 50 .txt files which consists of 10 comments each. The files were tested individually and the results were tallied for each tagged inappropriate expression, such as number of inappropriate expressions that must be tagged by the system, number of inappropriate expressions that were correctly tagged by the system based on the expert (TP), number of inappropriate expressions that were correctly tagged by the system but not the expert (FP), number of inappropriate expressions that were correctly tagged by expert but not the system (FN) and number of appropriate expressions that were correctly tagged by the expert and system (TN).

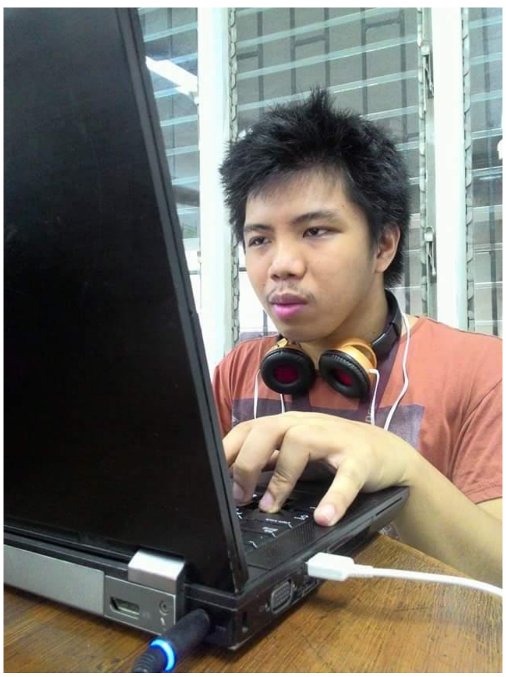
ISSUES AND CONCERNS

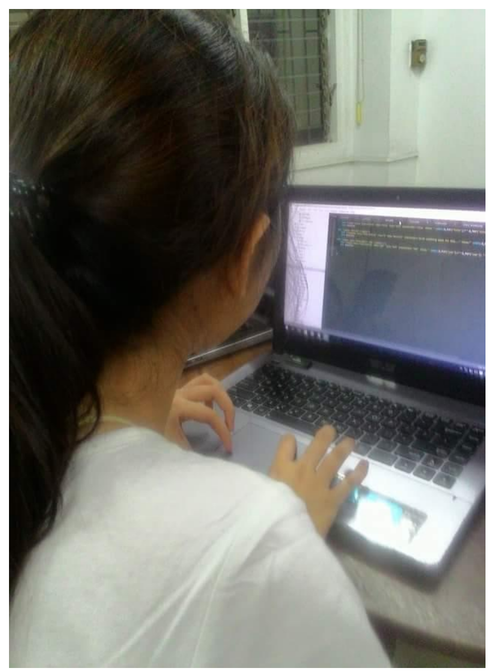
The following are the issues and concerns that were encountered during implementation:

1. Most of the comments found in the web are in shortcut format.
2. Time-consummation of gathering data.
3. Misuse of capitalization of letters.

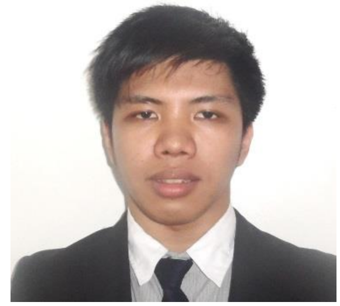








CURRICULUM VITAE



JOSHUA S. DAPITAN
Blk. 13 Lot 17 Jasmin St., BF Homes Phase III, Caloocan City
Contact Number: 9362570/09334091593
Email Address: joshuadapitan@gmail.com

EDUCATION

2012 – Present	**Polytechnic University of the Philippines** Bachelor of Science in Computer Science Sta. Mesa, Manila
2008 – 2012	**Graceland Christian College** Deparo, Caloocan, Metro Manila
2007 - 2008	**Holy Child Academy, North Fairview Annex** North Fairview, Quezon City, Metro Manila Salutatorian

PERSONAL INFORMATION

Name:	Joshua S. Dapitan
City Address:	Block 13 Lot 17 Jasmin St., BF Homes Phase III, Caloocan
Sex:	Male
Civil Status:	Single
Date of Birth:	January 31, 1996
Place of Birth:	Manila City
Height:	5’ 4”
Weight:	115 lbs
Religion:	Protestant
Citizenship:	Filipino
Contact Number:	9362570/09334091593
Parent’s Name	

Father:	Jude F. Dapitan
Occupation:	Geodetic Engineer
Mother:	Emma S. Dapitan
Occupation:	House Wife
Address:	Block 13 Lot 17 Jasmin St., BF Homes Phase III, Caloocan
Contact Number:	9362570

SKILLS

Computer Skills
Software: MS Word (Office Typing Jobs), MS PowerPoint (Slideshow Presentation), MS Excel (Spreadsheet), MS Publisher (Brochure Designing), Github (Collaboration/Repository), EasyWorship (Presentation), Movie Maker (Video Editing)
Programming Languages: C, Java, C#, Python, MATLAB, PHP, JavaScript, SQL
DBMS: MySQL, MS SQL Server
Operating Systems: Windows, Android

Languages Spoken
Filipino, English

SEMINARS ATTENDED

DLSU RESEARCH CONGRESS 2015
March 2-4, 2015
Henry Sy, Sr. Hall, De La Salle University, Manila
5th CCIS-DCS Computer Science Research Colloquium
February 17, 2015
Bulwagang Balagtas, NALLRC Building, PUP Manila (main campus)
Advance Topics in Computer Science Research
February 14, 2015
Claro M. Recto Hall, PUP Manila (main campus)
4th CCIS-DCS Computer Science Research Colloquium
February 18, 2014
Bulwagang Balagtas, NALLRC Building, PUP Manila (main campus)
Computer Research and Engineering Symposium 2016
January 21-23, 2016
Tanghalang PUP, PUP College of Communication Building, PUP Manila (main campus)

ANJANETTE R. LASALA
Blk.1 Lot 1 Junji St., Brgy. Kaligayahan, Novaliches, Quezon City
Contact Number: 09757342176
Email Address: anjlasala@gmail.com

EDUCATION

2012 – Present	**Polytechnic University of the Philippines** Bachelor of Science in Computer Science Sta. Mesa Manila
2008 – 2012	**Lagro High School** Flores de Mayo, Novaliches, Quezon City, Metro Manila Graduated
2007 – 2008	**Lagro Elementary School** Ascension Avenue, Novaliches, Quezon City Graduated
2002 - 2007	**Iemelif Learning Center Inc.** Greenfields-1, Kaligayahan, Novaliches, Quezon City

PERSONAL INFORMATION

Name:	Anjanette R. Lasala
City Address:	Blk. 1 Lot 1 Junji St. Kaligayahan, Novaliches, Quezon City
Provincial Address:	Lingayen, Pangasinan
Sex:	Female
Civil Status:	Single
Date of Birth:	October 22, 1995
Place of Birth:	Quezon City
Height:	5’3”
Weight:	49 kg
Religion:	Roman Catholic
Citizenship:	Filipino

Contact Number: 09757342176
Parent's Name
Father: Alfredo R. Lasala
Occupation: Factory worker
Mother: Jesselyn R. Lasala
Occupation: House Wife
Address: Blk. 1 Lot 1 Junji St. Kaligayahan, Novaliches, Quezon City
Contact Number: 09272106873

SKILLS

Computer Skills
Software: MS Word (Office Typing Jobs), MS PowerPoint (Slideshow Presentation), MS Excel (Spreadsheet), Adobe Photoshop
Hardware: Hardware Installation, Basic Networking
Database/Programming: C++, Java, C#, MySQL, MS Access, PHP, HTML, CSS
Other Skills
Encoding, Clerical and Office Jobs
Language Spoken
Filipino, English

SEMINARS ATTENDED

Computer Literacy Training Program
SY 2010-2011
Diliman Computer Technology Institute, Commonwealth Avenue, Diliman, Quezon City
Google Festival
December 10, 2012
Bulwagang Balagtas, Ninoy Aquino Library and Learning Resource Center, PUP, Sta. Mesa, Manila
ATCS: Symposium 2015: Advanced Topics in Computer Science
February 14, 2015
Claro M. Recto, PUP, Sta. Mesa, Manila
5th CCIS-DCS Computer Science Research Colloquium
February 17, 2015
Bulwagang Balagtas, Ninoy Aquino Library and Learning Resource Center, PUP, Sta. Mesa, Manila
DLSU Research Congress 2015
March 2-4, 2015
Henry Sy, Sr. Hall, De La Salle University, Manila