

**New Era University**

College of Computer Studies

Department of Computer Science

**DEEPTER: Stress Analysis for Twitter Post Using Natural Language Processing’s Sentiment Analysis**

by

**Sigua, Joenalyn Supan**

An Undergraduate Thesis Submitted to the College of Computer Studies in Partial Fulfillment of the Requirements for the Degree

**Bachelor of Science in Computer Science**

March 2018

i

**ABSTRACT**

**Title:**

**Author/s:**

**Adviser/s:**

**School:**

**Degree Program:**

**DEEPTER: Stress Analysis for Twitter Post**

**Using Natural Language Processing’s**

**Sentiment Analysis**

**Joenalyn S. Sigua**

**New Era University**

**Bachelor of Science in Computer Science**

It is of significant importance to detect and manage stress before it turns into severe problems. However, existing stress detection methods usually rely on psychological scales or physiological devices, making the detection complicated and costly. In this paper, the proponents explored to automatically analyze individuals' stress rate via social media. To achieve this goal, we import 7000 words in our database using MS SQL server. The system was developed using Natural Language Processing’s Sentiment Analysis and Support Vector Machine Algorithm for the classification of tweets that can detect at-risk users with an accuracy of 76%. This proved that the application has a high accuracy of classification. The purpose of this system is to serve as an automatic monitoring of stress and extra diagnostic tool for mental health professionals and to give every concerned individual the ability to analyze their beloved ones twitter posts who deserve warm help at an early stage. The web application also provides information about stress to help users to be aware of the stress in advance, and manage the stress before it becomes severe and starts causing health problems. The development of the application is discussed in this study.

***Keywords:*** *Stress Analysis, Twitter, Sentiment Analysis, and Machine Learning*

ii

**Chapter I**

**THE PROBLEM AND ITS BACKGROUND**

**INTRODUCTION**

Aided by the convenience and constant access provided by mobile devices, about 65% of adults and 92% of teens use social networking sites, a nearly tenfold jump in the past decade (Pew Research Center, 2015). The proponents rely on social networks to share ours or others’ daily activities with a wide audience. Besides, the proponents use social networks to disclose emotions and moods, happiness and unhappiness, since disclosure is intrinsically rewarding and can improve interpersonal intimacy. Inevitably, the language used in social media postings may signal feelings of exhaustion, sleeping problems, sweating, loss of appetite and difficulty concentrating that characterize the major stress. It is thus feasible to detect users’ psychological stress from social media data.

Twitter is a very popular social media platform in today’s digital era. Every second, 600 tweets are approximately published on Twitter (Yiping Li, 2015). Twitter implements limited character policy, it enables users to convey their message in a short version and as of September 26, 2017 Twitter has started testing 280-character tweets, doubling the previous character limit, in an effort to help users be more expressive (Casey Newton, 2017). This has encouraged the use of short forms to fit into the provided character limit. These short forms

1

involve the use of alphabets, digits and exclamations. This has been quite handy for today’s generation. Twitter also has the hashtag feature wherein any phrase or word written after a hashtag is considered as the topic of the tweet. Twitter trends are decided by the highest frequency of similar hashtag phrases all around Twitter. The most discussed topics become trends for the day. Twitter also offers regional trends, mainly country-wise.

Since the boom in the E-commerce section, social media platforms are crucial in customer feedback. Sentiment analysis of a product is a major field of study in this culture. But, along with product sentiment and opinions, human emotions can also be analyzed by tweet contents. Twitter offers a quick self-expression format, which is unique because of features like retweeting, hashtag trending, limited characters etc. Twitter gives a chance for quick updates to everyone. The Twitter behavior pattern of a person can be very useful to study human emotions and changing form of self-expression, along with all its effects on other aspects of social life. For such study, Teenage Twitter users prove to be a good subject. Tweets from such users are mostly spontaneous, unfiltered thoughts expressed without least possible regard for its consequences. This can help in identifying correct emotions of a person at that point. Teenage users are also constantly active, which makes a large chunk of continuous data to analyze.

In psychology, stress is our bodies’ response to any kind of frustrations, demand or threat (Jeanne Segal et al., 2018) which fills up our modern life. Thereby, stress is so commonplace that it has become a way of life. Some stress

2

is normal and even useful. For example, it can help us win a race or finish an important job on time. But when stress becomes overwhelming or lasts too long, it can damage our health, mood, relationships, and quality of life. Also, it can increase the risk of strokes, ulcers, and mental disorders. It is hence of importance to recognize the signs of stress in advance, which enables us to take proactive care to reduce the harmful effects of stress. It is worth mentioning that few efforts thus far have been dedicated to stress detection by harvesting social media, except the followings. (Lin et al., 2014) and (Xue et al., 2013) proposed novel methods to detect tweet-level stress and user-level stress over a short time window respectively. Despite great success, they only classified samples into stressed or non-stressed categories, which is incapable of measuring the exact stressor and stress level. Stress is indeed much more complicated, which is composed of two key factors: stressor and stress level. Quite literally, stressor, comprising of stressor event and stressor subject, triggers stress; meanwhile, different stressors incur different stress levels. For instance, a layoff usually makes people more stressed in comparison to a project deadline. In addition, stressor events happening to other subjects can also be someone’s stress trigger, but may have different effects. For instance, “my friend’s father just passed away”. Moreover, stress level depends on the stressor, and is measurable using various psychological stress scales, despite that stress is often thought of as a subjective experience. As claimed by psychologists, detecting stressor and measuring the stress levels are of essential importance to proactive care.

3

In this paper, the proponents work towards measuring stress via social media data. The goal of this research is to offer social workers the ability to access potential depressive people who deserve warm help at an early stage, as well parents of teenagers, and concerned friends alike, the ability to be attentive to their beloved one’s psychological status, without invasion of privacy, to ensure the well being of their friends and relatives. Whilst the work will not claim to replace or represent any type of professional mental health care, it will be useful as a tool to monitor a specific threshold of decline in the perspective of the user. The sentiment result is based on the analysis of texts from the user’s Twitter account. The concept of the work is to rate the user’s level of stress via social media with the use Natural Language Processing’s Sentiment Analysis Techniques with incorporation of Machine Learning Algorithm.

**Background of the Study**

Stress is becoming a threat to people’s health nowadays. With the rapid pace of life, more and more people are feeling stressed. According to a worldwide survey reported by Newbusiness in 2015, over half of the population have experienced an appreciable rise in stress over the last two years. Though stress itself is non-clinical and common in our life, excessive and chronic stress can be rather harmful to people’s physical and mental health. According to existing research works, long-term stress has been found to be related to many diseases, e.g., clinical depressions, insomnia etc.. Moreover, according to the Chinese Center for Disease Control and Prevention, suicide has become the top

4

cause of death among youth, and excessive stress is considered to be a major factor of suicide. All these reveal that the rapid increase of stress has become a great challenge to human health and life quality. Thus, there is significant importance to detect stress before it turns into severe problems. Traditional psychological stress detection is mainly based on face-to face interviews, self-report questionnaires or wearable sensors. However, traditional methods are actually reactive, which are usually labor-consuming, time-costing and hysteretic. Are there any timely and proactive methods for stress detection?

The rise of social media is changing people’s life, as well as research in healthcare and wellness**.** With the development of social networks like Twitter and Facebook and more people are willing to share their daily events and moods, and interact with friends through the social networks. As these social media data timely reflect users’ real-life states and emotions in a timely manner, it offers new opportunities for representing, measuring, modeling, and mining user behavior patterns through the large-scale social networks, and such social information can find its theoretical basis in psychology research. Twitter, in conjunction with natural language processing and machine learning, can be leveraged to support the analysis of very large data sets for population-level mental health research (Conway and O’Connor, 2016). Tools for analyzing social data are already successfully being used in other fields like consumer behavior, education, and [crime prediction.](http://www.nytimes.com/roomfordebate/2015/11/18/can-predictive-policing-be-ethical-and-effective/social-media-will-help-predict-crime) The purpose of this study is to be able to give

5

information and awareness to people about stress with the use of Natural Language Processing’s Technology.

**Objectives of the Study**

The main objective of the study is to design and develop a Web Based Application that will analyze the signs of stress on twitter posts using Natural Language Processing’s Sentiment Analysis and a Supervised Machine Learning Algorithm. The objectives help the researchers to determine the application is useful and needed are the following:

**Specific Objectives**

1. To extract sentiments from Twitter using different extracting techniques.
2. To implement algorithm and techniques that will analyze signs of stress from twitter data.
3. To evaluate the accuracy of the developed application in terms of precision, f-measure, recall.
4. To evaluate the usability of the system.

**Significance of the Study**

In this part of the study, it provides information by the researcher on how the study will contribute to the beneficiaries. Specifically, the proposed study will be significant for the following:

6

**Teenagers.** More and more teenagers today are suffering with adolescentstress from different aspects: academic future, self-cognition, interpersonal, and affection. Long-lasting stress may lead to anxiety, withdrawal, aggression, or poor coping skills such as drug and alcohol use, threatening teenagers’ health and development. This system will help teenagers to be aware of the stress in advance, and manage the stress before it becomes severe and starts causing health problems.

**Health Care Professionals.** This web application could be a monitoringtool for health care professionals to be able to track their patients’ emotional states via twitter or to recognize anxiety or systemic stressors of populations (e.g. different student groups on campus) via twitter data.

**Concerned Individuals.** This tool offers social workers the ability to accesspotential depressive people who deserve warm help at an early stage, as well parents of teenagers, and concerned friends alike, the ability to be attentive to their beloved one’s psychological status, without invasion of privacy, to ensure the well being of their friends and relatives.

**Commercial Agencies.** The system can also help commercial agenciesto gauge the sentiment of buyers or to facilitate targeted product advertisement.

**Researchers.** This shall address as the references of other upcomingprogram developer of the implemented software.

**Students.** Who wants to conduct and continue the study for furtherclarifications.

7

**Scope and Limitations**

Social media has recently emerged as the head technique to disperse data on the online network. Through these online systems, countless people convey their considerations, individual encounters, and social standards. Individuals who are not that alright with their specialists tend to post their feelings and emotions on social media. This data could be utilized to help these patients who are suffering psychological stress. So the scope of this project is much broader as stress is the root cause of many health problems and mental diseases. Anticipated to become one of the prominent problems in recent years in our modern society. The project is to analyze stress among individuals using their bundle of tweets The project is a great contribution to the health of society. As it provides a tool for the psychiatrists and psychologists to help them understand their patients a bit more. It could help them study even the history of their patients using some old tweets. A lot of benefits could be deduced from this project. This project has made the diagnosis easier as the patient doesn’t have to go through some harsh tests for diagnosis.

But the present study also has several notable limitations. First, the code is never as smart as a human. The project does not handle the sarcasm as there is a possibility where the patient has tweeted something negative that has a hidden humorous meaning. Irony, diversion and different nuances of human discourse, similar to how the emoji can change the tone of a generally negative articulation. Spam-loaded discussions in online networking that strike individuals as inauthentic. False negatives, where the product sees a negative word like "crap". However doesn't understand it's positivity in the general setting as"Holy

8

crap! I love this!” and the tweeted phrase “You’re killing it!” may either mean “You’re doing great!” or “You’re a terrible gardener!”. Cultural contrasts, where a few people from a few nations may be pretty much unreserved in their utilization of dialect. But as the sentimental analysis deals with the probability of negativity and positivity of the unigrams and bigrams so there might be a chance that a sentence as a whole makes a different sense. So there are times when one can question the accuracy of the results of this project as the tweet doesn’t show all the hidden circumstances of that individual. The system will be applied in all the twitter user’s with public account and is limited to English words only.

9

**CHAPTER II**

**REVIEW OF RELATED LITERATURE**

In this chapter, the proponents provide a literature review of studies related to mental illness in social media, including an overview of methodologies used by existing systems.

**Stress Analysis**

Computer-aided detection, analysis, and application of emotion, especially in social networks, have drawn much attention in recent years (Coppersmith et al., 2014), (Zhao et al. 2014), (Lin et al., 2014), (Thelwall et al., 2015), (Zhang et al., 2015). Relationships between psychological stress and personality traits can be an interesting issue to consider (Farnadi et al., 2016), (Golbeck et al., 2014), (Verhoeve et al., 2016). For example, (Bogomolov et al., 2014) providing evidence that daily stress can be reliably recognized based on behavioral metrics from users mobile phone activity. Many studies on social media based emotion analysis are on the tweet level, using text-based linguistic features and classic classification approaches. (Ke Xu et al., 2013) proposed a system called MoodLens to perform emotion analysis on the Chinese micro-blog platform Weibo, classifying the emotion categories into four types, i.e., angry, disgusting, joyful, and sad. (Xue et al., 2014) studied the emotion propagation problem in social networks, and found that anger has a stronger correlation between different users than joy, indicating that negative emotions could spread more

10

quickly and broadly in the network. As stress is generally considered as a negative emotion, this conclusion can help us in combining the social influence of users for stress detection. However, these works to mainly leverage the textual contents in social networks.

In reality, data in social networks are usually composed of sequential and interconnected items from diverse sources and modalities, making it be actually cross-media data. Researchers are trying to leverage pervasive devices like personal computers and mobile phones for routine stress detection. (Frauendorfer et al., 2013) proposed StressSense to unobtrusively recognize stress from human voice using smart phones. (Paredes & Canny, 2013) investigated the initial lab evidence of the use of a computer mouse in the detection of stress. However, such applications rely on collecting one’s real-life data, which is easy to trigger antipathy. It makes stress detection invasive to normal life, and can't be used widely in more people.

Stress monitoring and prediction techniques can be divided into three main kinds according to the stress detection methods being used (N. Sharma et al., 2015). The first kind uses subjective questionnaires or individual/group meetings with psychologists to analyze users’ stress situations. This kind of methods needs high cooperation of users and sometimes relies on people’s ability to recall their experiences. Observing human’s physiological and physical signals change with the variation of psychological stress status, in the recent ten years, many researchers used various sensors to objectively monitor the changes and predict the trends of physiological signals and physical signals. For

11

example, (J. Bakker et al., 2014) used physiological symptoms of GSR signals and accelerometer data about movements to measure user’s stress level, and discovered correlations between the changes in GSR and the characteristics of the events using associative classification in data mining. Compared with these body contact and invasive stress measurements, the open micro-blog arises as another low-cost sensing channel to obtain people’s self-expressed contents and behaviors, from which some emotional signals could be captured and analyzed. For example, (Park et al., 2013) evaluated whether people are in the risk of depression by analyzing their twitting behaviors. Recently, (Xue et al., 2014) investigated a number of teens’ typical tweeting behaviors that may reveal adolescent stress, and applied five classifiers to teens’ stress detection.

**Twitter**

Due to the widespread adoption of social media and the availability of large-scale data from social media, approaches that use such data for depression screening are receiving increased attention from researchers. (Moreno et al., 2011) shows that college students display symptoms consistent with depression on Facebook, a popular social networking service. (Park et al., 2012) analyze differences between Twitter users with and without depression by analyzing their activities. Both (De Choudhry et al., 2013) and (Horovitz et al., 2013) used crowdsourcing to identify Twitter users reported to be in depression, according to the standard psychometric measures. (De Choudhry et al., 2013) have used behavioral characteristics identified under engagement, egocentric social graph, depression language, emotion and linguistic style to determine the cause of

12

major depressive disorder symptoms. With an accuracy of 70% and a higher precision of 0.74 for the depression class, the authors managed to identify vulnerable individuals to have depression before the start of their major depressive disorder. As contributors to the mentioned accuracy, reduced social activity, increased negative affect, clustered social network of the individual, raised interpersonal and medical fears and increased expression in religious involvement, were identified as strong indications leading towards major depressive disorder. Even though the impacts of the ego network among individuals who are vulnerable to depression are being identified as considerably small, further analysis has shown that such networks are tightly clustered and close-knitted. Similar to De Choudhury et al. (2013), Tsugawa et al. (2015) have demonstrated the effectiveness of using social media platforms among Japanese Twitter users in order to recognize depression.

In addition to using Twitter user's activity history, a web based questionnaire was used to collect ground truth data in predicting the existence of depression for the Twitter users. In addition to the features used by De Choudhury et al. (2013), (Tsugawa et al.) used bag of words and word frequencies to identify the ratio of tweet topics. Even though subtle changes in behavioral features were identified between their research and the one done by De Choudry et al. which could be due to cultural aspects, the authors have identified similar analytical patterns for the use of negative words, posting frequency, retweet rate, and the tweets containing URL's. Feature engineering using Twitter user activity positively contributed towards a classification accuracy

13

of 69%, with 0.64 precision, and 0.43 recall using support vector machine (SVM) classification. In (Park et al., 2013) a similar analysis is performed by analyzing data from Facebook. Using multiple regression analysis, (Tsugawa et al., 2013) show that the frequencies of the word usage on Twitter are useful features for recognizing depression among the users. The main objective of such research is to clarify which features that can obtained from user activity are useful for estimating the severity of depression. (De Choudhury et al., 2014) are pioneering in demonstrating the estimation accuracy that could be achieved by using activities on Twitter to predict depression among the users. In their study, De Choudhury et al. obtained training data for machine learning by appealing to large numbers of people for help (popularly known as crowdsourcing).

Models that could be used to predict risk of depression were identified from several features obtained from the records of user activity on Twitter by using SVM. Experimental results show that depression can be recognized among users with an accuracy of approximately 70%. Such approaches are applied to prediction of postpartum depression from Facebook and Twitter data (De Choudhury et al., 2013, p.3267). De Choudhury et al. have also proposed a method for estimating the depressive tendencies of populations by a similar approach (De Choudhury et al., 2013, p. 47). Twitter data are also suggested to be effective for estimating health-related statistics (Culotta et al., 2014)(Paul et al., 2011). Research into using social media data for recognizing depression in individuals is just beginning, and so whenever possible the effectiveness of such approaches should be validated for several datasets. Our study builds on the

14

mentioned prior work and contributes to enhancing methods for predicting depression risk from objective information, particularly large-scale data from social media.

**Sentiment Analysis**

Tejwani, who conducted a survey on sentiment analysis, states that ”Sentiment analysis is also known as opinion mining in which text analysis, natural language processing techniques are used to identify sentiments ”(Tejwani, 2014).

This survey covered different techniques used in sentiment analysis and opinion mining. It states, “Sentiment classification classifies text as per different opinions towards certain object. Feature based sentiment classification considers the opinions on features of certain objects” (Tejwani, 2014). The sentiment analysis task is classified into three different levels: document level, sentence level, and feature bases techniques i.e. aspect level and comparative analysis (Liu, 2010). Sentiment analysis on micro-blog social media like Twitter have been conducted for various applications. Previous work on sentiment analysis investigated users’ tweeting contents, hashtags, emoticons, tweeting/ retweeting/mentioning behaviors (J. Read, 2015). (Bollen et al., 2014) performed sentiment analysis, research of all public Twitter posts over a period of four months. They used a syntactic term based approach to measure the sentiment of tweets via a psychometric instrument called

15

POMS. This approach showed that supervised learning is not the only viable way to perform sentiment analysis of Twitter. (Bollen and Pepe, 2014) analyzed the mood expressed in 10,741 emails in the future. To score the moods of the emails, they used an extended Profile of Mood States metric. They extended POMS original 65 adjectives with WordNet 3.0 synonyms. The extended list of POMS words were then stemmed using the Porter Stemmer. They scored the words Pak and Paroubek used Twitter as a corpus for sentiment analysis and opinion mining. Their method studied the POS tag distribution differences between positive, negative, and neutral tweets. A multinomial Naive Bayes classifier based on POS tags and n-grams were used. They concluded that a Twitter user’s emotion is reflected in the syntactic structures of their tweets. (Lu et al., 2015) created a framework which automatically constructs a context dependent sentiment lexicon. An unambiguous, gold standard sentiment lexicon is used as the basis. The polarity of these sentiments is propagated into another aspect-word pairs through language clues, a synonym antonym and overall review ratings. In conclusion they found the framework could successfully learn new aspect dependent sentiments.

**Machine Learning**

The use of Machine Learning (ML) is relatively new in the field of clinical psychology, specifically in its use to predict suicide risk. ML implements algorithms to classify complex problems.

16

Several efforts have attempted to automatically detect depression in social media content using machine learning. The work presented by (Lin et al., 2014) identifies user psychological stress in tweets. Features such as emotion words, smileys, tweet mentions, replies, and posting frequency were obtained from single tweets, and from all user’s tweets. The best performance was obtained by a four layer Deep Neural Network (DNN). (Preotiuc-Pietro et al.,2015) employed user metadata and textual features from the corpus provided by the CLPsych 2015 Shared Task to develop a linear classifier to predict users 33 having either one of the mental illnesses. They have used the bag-of-words approach to aggregate word counts, topics derived from clustering methods and metadata (e.g., followers, followees, age, gender) from the users Twitter profile as the main feature categories. With the use of logistic regression and linear SVM in an ensemble of classifiers, the authors managed to obtain an average precision above 0.800 for all the three tasks and with a maximum score of 0.867 for differentiating users in the control. (De Choudhury et al., 2013) presented a study on predicting depression from tweets by analyzing over 2 million posts by 476 users. The best performance was obtained with an SVM classifier and a set of behavioral features, such as occurrence of pronouns, usage of swearing and depression terms, tweet replies, as well as posting time and frequency. (Resnik et al., 2015), who was ranked first in the (CLPsych,2015) shared task created 16

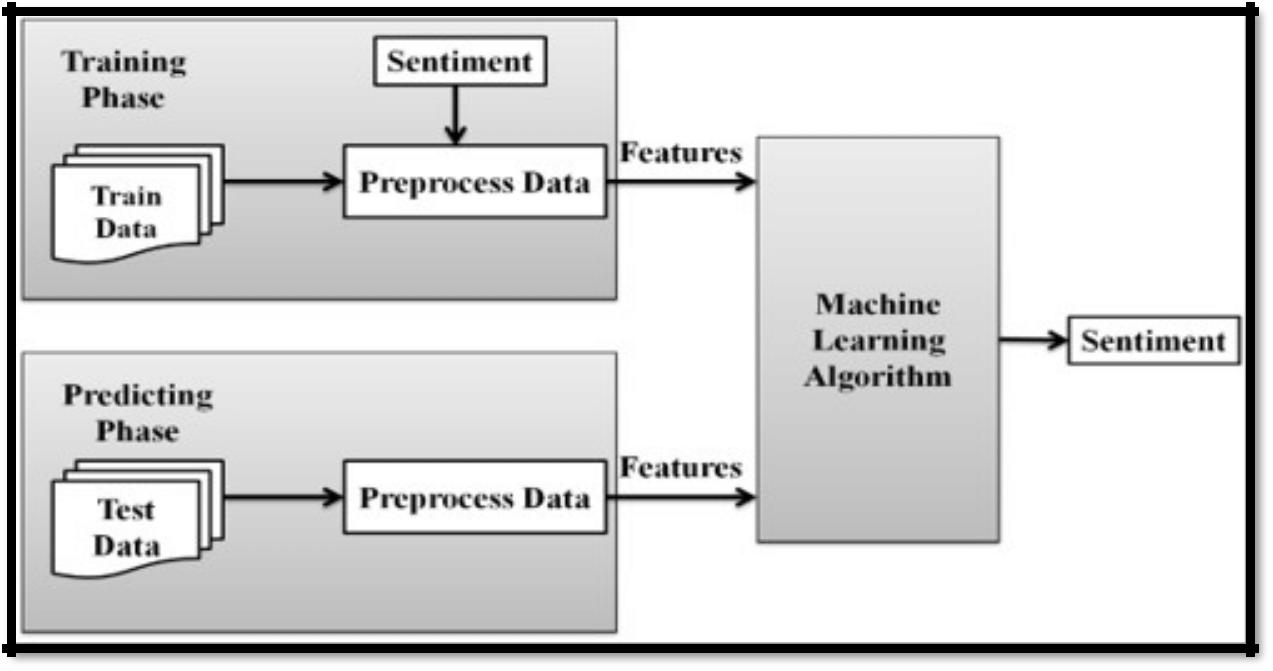
17

systems based on features derived using supervised LDA, supervised Anchor (for topic modeling), lexical TF-IDF and a combination of all. They highlighted the importance of TF-IDF features, which has made a significant contribution in obtaining higher accuracies used by itself as well as in combination with supervised topic modeling. The authors used SVM classifiers with linear or RBF kernel. They grouped the dataset into weeks 32 to avoid the issue of the data set being too large or small. Finally, the dataset was pre-processed using basic pre-processing methods (e.g., remove emoticon character codes, removing stop words, lemmatizing). With 80% training and 20% testing data split, they managed to obtain the best results for all three tasks using an SVM classifier with linear kernel and using the big vocabulary for computing the models using all the features. The classifier obtained an average precision above 0.80 for all the three tasks and a maximum precision of 0.893 for differentiating PTSD users from the control group. ( Pang et al., 2007) investigated three machine learning methods to produce automated classifiers to generate the class labels for movie reviews. They tested on Naïve Bayes, Maximum Entropy and Support Vector Machine, and evaluated the contribution of different features including unigrams, bigrams, adjectives and POS-tags. Their experimental results found that the SVM classifier with unigram presence features outperformed other competitors.

18

**Theoretical Framework**

The theories used in this study, the researcher is the beneficiary in classifying to understand and indicate the process of proposed study.



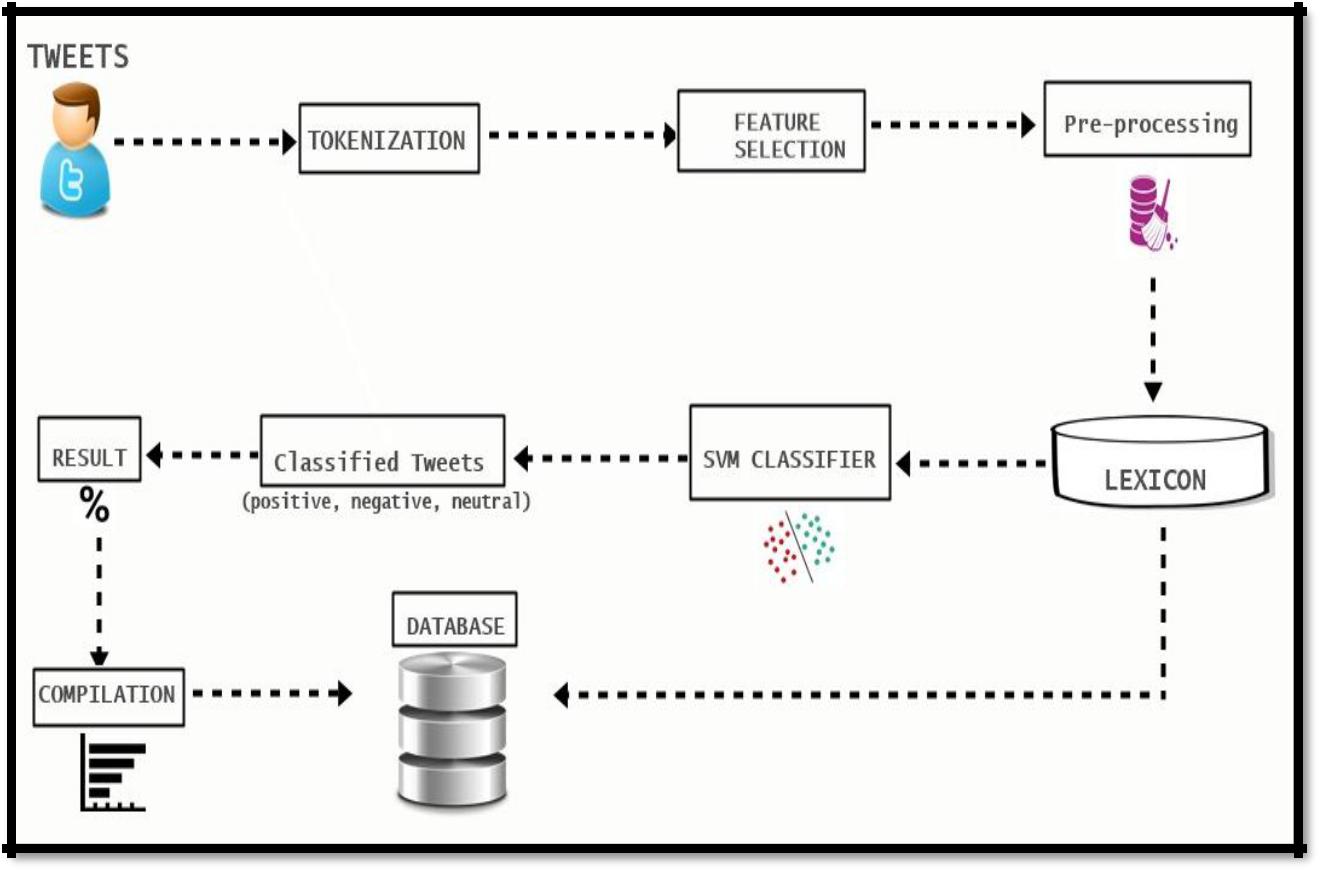
**Figure 2.1: Sentiment Analysis Phase (Arockia et al., 2016)**

The sentiment analysis task is divided into two phases: training phase and predicting phase. The training phase involves loading train data (manually classified reviews) which are used to train the machine learning algorithm. The predicting phase involves loading test data (unclassified reviews) for which sentiments need to be determined. Both train and test data are pre-processed to obtain meaningful features before sentiment is obtained.

19

**Conceptual Framework**

This study is to give ideas in illustration and representation on how the study works, how it is processed and its flow.



**Figure 2.2: Conceptual Framework**

In this framework, the first phase is getting the recent tweets of the user with the use of Twitter Application Interface (API). Once the twitter data is fetched, it will be delivered to the tokenization phase to count the polarity words inside the tweets. Tokenization means to break a sentence into words. The next phase is to capture features that describe the emotion expressed by each tweet. Feature selection plays an important part in the effectiveness of the classification process. After selecting the features, tweets are then pre-processed to mitigate misspellings used in Twitter.

20

The Lexicon contains all the words and polarity score that will be used to compute the sentiment score of the tweets. The text/tweets will be classified using the supervised machine learning algorithm called Support Vector Machine that will result to categorization of tweets (Positive, Negative, Neutral). After the classification process, the result will be the percentage rate of the recent tweets analyzed; the system will also compile all the analysis results of the recent users.

**Definition of terms**

**Confusion Matrix.** Type of evaluation technique to get the accuracy or

percentage of a system.

**Dataset or Trained Dataset.** Is a collection of pre-processed data to usein faster pre-processing of test data.

**Lexicon-based sentiment analysis**. Application of a lexicon is one of thetwo mainapproaches to sentiment analysis and it involves calculating the sentiment from the semantic orientation of word or phrases that occur in a text.

**Natural Language Processing(NLP).** NLP can analyze languagepatterns to understand text. One of the most compelling ways NLP offers valuable intelligence is by tracking sentiment the tone of a written message (tweet, Facebook update, etc.) and tag that text as positive, negative or neutral.

21

**N-gram model.** Can be used to store this spatial information within the

text.

**Stress.** Psychological stress is becoming a threat to people’s healthnowadays. With the rapid pace of life, more and more people are feeling stressed. Stress can be caused by a number of factors such as peer pressure, family, work, parental pressure, exam, etc.

**Social Media.** Social media has become an integral part of everyday lifewhere many people have started sharing their day-to-day activities. With its rapid growth among different demographics, social media can be a significant contributor to the process of mental disorder detection and prevent serious consequences like suicide.

**Supervised machine learning.** A model is trained on a pre-defined set oflabeled training examples; then the model can be applied on new data in order to make predictions. Supervised learning problems are also referred to as classification problems.

**Support Vector Machine Algorithm.** A high performing machine learningalgorithm used for text classification.

**Twitter.** Provides a rich source for studying people’s mood. A tool used formeasuring and predicting stress in individuals.

**Twitter API**. A public API which enables programmatic collection of tweetsas they occur, filtered by specific criteria.

22

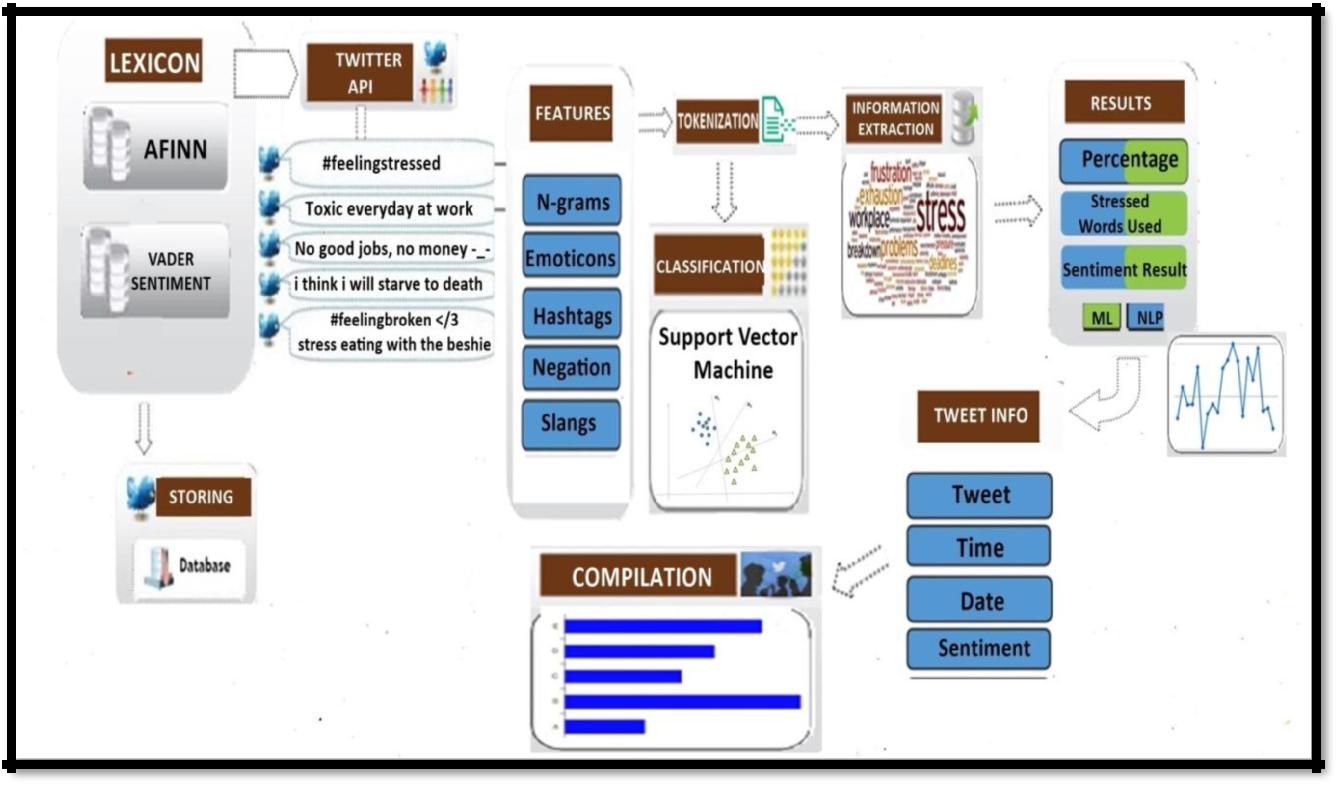
**Chapter III**

**Methodology**

This chapter discusses all the methods and techniques the researchers applied in the development of the system. It will provide insights on choices and decisions made by the researchers in order to conduct this study. Project Design, Operations Tools and Procedures, Project Evaluation Techniques and tools are also tackled in detail in this section of the study. .

**Project Design**

This part shows the project design part and the flow of the system.



**Figure 3.1: Project Design**

23

The project design will have two sides as it is a web based application. These two sides are front-end and back-end. The front-end of this project will be in AngularJS while the C# based RESTful web services will serve as back-end. These front-end and back-end sides can also be referred as client-side and server-side development respectively.

**Phase 1: Getting Twitter data.** On the client-side the first page willprompt the user to enter a username for which he/she want to analyze the tweets. The user can also specify the number of previous tweets he/she want to analyze. If the user does not provide any number then by default the last 20 tweets will be picked for analysis. Once the user has provided all the necessary information the client-side will request the server-side or RESTful web services to fetch the required tweets. On receiving the request from client-side, the web-service will fetch the required number of tweets from twitter using the Twitter API. Once the data is fetched, all these tweets and their data will be sent back to the client-side for the analysis.

**Phase 2: Analysis.** After having all the required data for analysis, Tweetswill be tokenized and then each tweet will be individually analyzed for negative, positive and neutral words. Every tweet will have a positive, negative and comparative score after the analysis. If the collective score of the tweets is below 25% then it means the user has positive posts, and if the collective score of

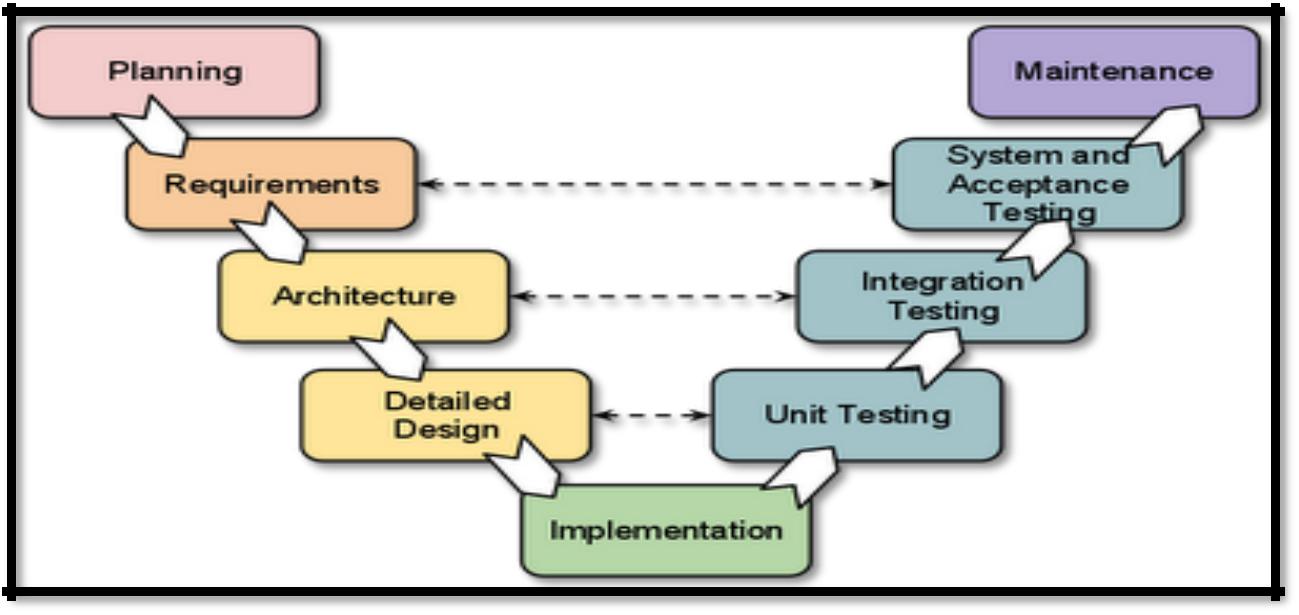
24

tweets is between 25-60% that means user has normal posts. But if the collective score of tweets exceeds 60% then that user has an unusual negative score which points out to the signs of stress.

**Phase 3: Information and Result.** The results will be shown back to theuser in the results page, along with the collection of negative and positive words used in the list of tweets. The results will also show the time and dates on which these tweets were published.

**Project Development**

This methodology part further discusses the phases that are used for the development of the study.



**Figure 3.2: V-model SDLC**

For the development of the website application project, the proponent utilized the V-Model methodology of SDLC which aims to produce the best quality software by means of testing and verification that takes place in the early

25

stages of the project to make sure that the conversion to the next stage is possible.

The development cycle starts in initial planning, the planning will be by defining the problem that every individuals are facing nowadays which is stress as teenagers, teachers, successful people even successful people, parents are all suffering from stress due under the rapid pace of life due to the rapid development of modern society. After defining the problem we then decides what is the best domain to use, we chose twitter as our social media domain because Twitter posts/Tweets are informal, with 500 million tweets per day(Halder, 2017), has a wide array of topics and large vocabulary, contains meta information like retweets, date and time, has special strings such as hashtags, emoticons, internet slangs, and stress sentiments which are the perfect features for having an accurate analysis result. After deciding what social media domain to use, we then planned on what algorithms, and techniques that we can use for the development of the application which is Natural Language Processing’s Sentiment Analysis, Information Extraction with the used of Support Vector Machine Algorithm which is the best approach for text categorization.

After the initial planning, next phase we used is the requirement analysis. For the requirement analysis we then gathered as much information as possible about the details of our desired software. For gathering information we then decide what features to include on our site plan and what lexicon or datasets to use. After having all the clear and detailed product requirements, it’s time to

26

design the complete system. The design of the system architecture along with its data flow representation should be clearly defined in the design stage, so in this stage we planned all the programming languages to use, the design will have two sides as it is a web based application. These two sides are front-end and back-end. The front-end of this project will be in AngularJS while the C# based RESTful web services will serve as back-end, and for the database we will use MS SQL database for the storing of our data sets. Actual development starts in the implementation stage where the developer follows the project design and specified programming tools to be used with respect to the development of the system application like which programming language will be used for coding.

For the Testing stage we used the procedural testing methodology, unit testing, system testing and integration testing that helps in the code coverage, debugging, and documentation of the project. It utilizes strict rules for each test level so that each component and function the documentation depicts is tested. A successful testing and meeting client expectations and requirement means the system is ready for the evaluation before the deployment of the system. Though the system was successfully tested and deployed to the client, maintenance of the system must be observed for monitoring and possible upgrading of the system.

27

**Testing and Operation Procedures**

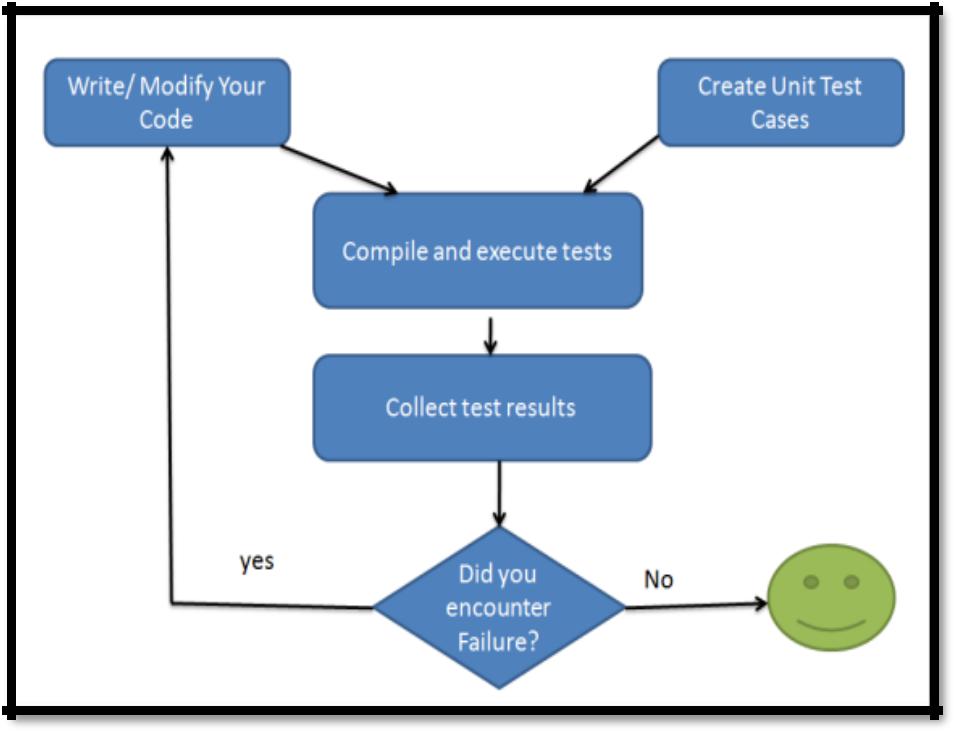
There is no software that is flawless. But one can make it reasonably flawless by performing different tests on it. Testing helps very much in debugging, code coverage and documentation of the project. For this project Procedural Testing Method is used. It utilizes strict rules for each test level so that each component and function the documentation depicts is tested. These test paths are particular advances, keystroke by keystroke, that achieve the outcome the test case characterizes.

These levels of testing helped the proponents to identify and fix the bugs before it becomes operational.

**Unit Level**

At this level every unit is tested individually and then the dependent ones are tested according to their identification and command of interfaces with all other units. For this testing we utilized a software testing method called White Box Testing. The functions, classes and methods of the web application were tested to detect defects in logic, data by testing individual modules of the system.

28



**Figure 3.3: Unit Testing Flowchart**

The following are the defects that occurred at unit testing:

* + Incorrect data type
* Event handling
* Variable declared but never used

**System Level**

At this level all the functions and components of a single unit are tested at a time. Interactions between one another and their compatibility with each other are tested as well. The objective of system testing is to verify whether all requirements are met as earlier defined.

The following are the defects that occurred at system testing

29

* Calculation of data is incorrect or inconsistent.
* Spelling errors in pages

**Integration Level**

At this level of testing includes testing the framework's expected design all in all and incorporate programming, firmware, as well as peripherals. All the units involved in the project are tested in accordance with the other units. During integration testing the proponents checked how one or more units interact with each other and produce output for various scenarios.

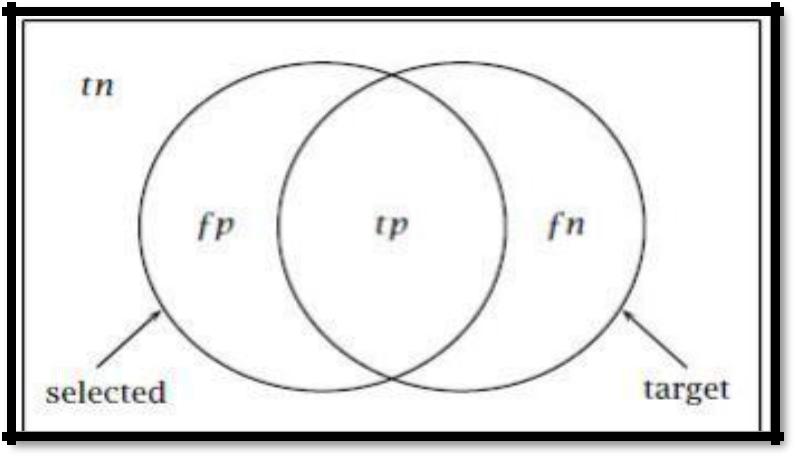
In this form of testing a lot of defects related to functional, requirement and performance levels were detected:

* Poor network configuration
* Long Load time
* Database configuration

**Project Evaluation Techniques**

The performance of the system is visualized with the use of confusion

matrix.



30

**Figure 3.4: diagram for evaluation definition**

This diagram is used to illustrate how the different metrics are defined in information retrieval (IR). In this diagram, there is the total data (the test data), let's say a collection of documents, represented by the rectangle. Let's say the IR system wants to retrieve a set of documents that fulfill some requirement. These set of documents are the 'target' (tp+fn). But the system actually retrieves the set of documents that are in the circle named 'selected' (fp+tp) in the diagram. tp stands for true positives, fp stands for false negatives and fn stands for false negatives. Those that are neither in the 'selected' nor in the targeted are true negatives (tn). From here, accuracy, precision, recall and F measure are computed and defined as follows:

**Precision**.“How many of the users we identified as stressed are actuallystressed?”

*Formula* : =

**Recall*.*** *“Out of all of the stressed users, how many did we properly detect?”*

*Formula =* =

**F-measure***.**Weighted harmonic mean of the precision and recall of the test*.

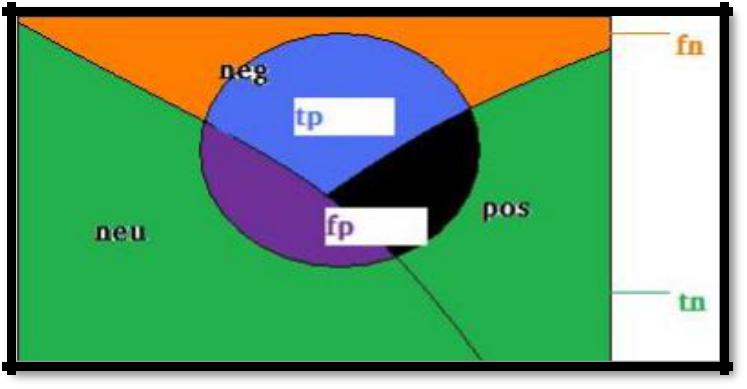
*Formula* =

**Accuracy**

31

*Formula:*

Accuracy measures the proportion of documents that are correctly obtained, precision measures the exactness, that is, the percentage of selected documents that are the targeted documents, and recall measures the completeness that is the proportion of targeted documents that the system selected. F measure is geometric mean of precision and recall. The measures of precision and recall are important performance measures when the dataset (set of documents) can be seen as relevant (the target) and irrelevant (outside the target). They are also useful measures to talk about the performance of a system when tn is so big. There can be a trade between precision and recall since one may be increased at the expense of the other. Depending on the nature of a NLP task, one may look for higher precision, recall, F measure or accuracy or some mixture of them all.



**Figure 3.5: diagram for evaluation definition on classification**

Now how do these performance measures apply in classification task, especially in a three-way or two-way classification that we have here. Again let's see a diagram with three classes neutral, negative and positive. The rectangle

32

(the whole test data) is divided into neg, pos, and neu. Let's say the neg class is now the target. The circle shows the selected. The circle does not exactly match the neg; it misses some part of the neg, and takes some parts from neu and pos. The part that is missed from neg is fn and the parts that are taken from neu and pos are *f p*. The parts that are outside the selected (the circle) and neg are tn. This is when we target the negative class. But, in three-way classification, there are three targeted classes. Thus there will be similar results for positive and neutral classes with their own *tp, fn, f p* and *tn*. The definitions of precision, recall, accuracy and F measure will be exactly like we saw above in terms of the *tp, fp,* *fn,* and *tn*. For a three-way classification, the results can be given in the form of aconfusion matrix which is also called contingency table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **a** | **b** | **c** | **<-classified as** |
|  |  |  |  |  |
|  | **a:a** | a:b | a:c | Negative |
|  |  |  |  |  |
|  | **b:a** | b:b | b:c | Positive |
|  |  |  |  |  |
|  | **c:a** | c:b | c:c | Neutral |
|  |  |  |  |  |
|  |  |  | **Table 3.1 Confusion Matrix** |  |

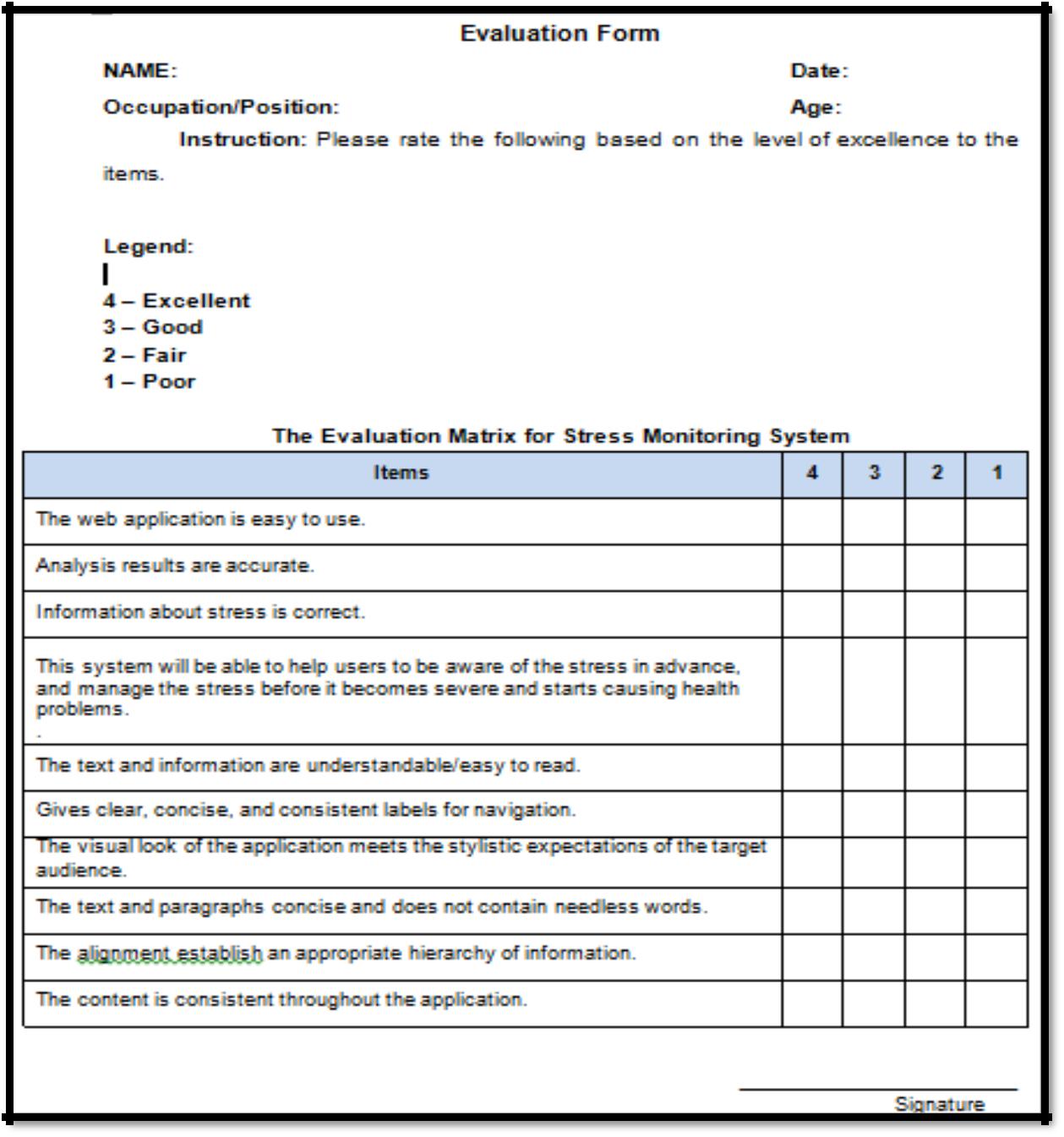
The confusion matrix shows the actual and predicted classes. The rows show the actual class. To make it clear and to relate it with the diagram above, let's see for the negative class. There are *a : a + a : b + a : c* total number of actual negatives. Out of them, the classifier got only *a : a* correctly. It missed *a : b* and *a : c.* Instead it classified b : a and c : a as negatives, which actually are positive and neutral respectively. *a : a* is the *tp* in the diagram, a : b and *a : c*

33

constitute *fn*, and b : a and *c : a* constitute *f p*. The rest, that is *b : b, c : c, b : c, c*

* *b* and *c : c* constitute tn. Similarly, *b : b* is the number of positives that theclassifer got right, and *c : c* is the number of neutrals that the classifier got right.

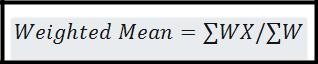
**Usability**



**Figure 3.6 : Evaluation Form**

34

This evaluation form will be used by the proponents to get the result for the usability of the system. The evaluators will be Registered Psychometricians who has a great knowledge about our study that is related to Psychology. The respondents will also be one of the beneficiaries of the application.



∑ = the sum of W = the weights X = the value

This is the formula that will be used on calculating the weighted mean on the survey of the system. The summation of all the values in your weights and divide it with the summation of the number of weights in your data set.

|  |  |  |
| --- | --- | --- |
| **Scale** | **Grade** |  |
|  |  |  |
| **3.51--> 4.0** | **Excellent** |  |
|  |  |  |
| **2.51 --> 3.50** | **Good** |  |
|  |  |  |
| **1.51 --> 2.50** | **Fair** |  |
|  |  |  |
| **1.0 --> 1.50** | **Poor** |  |

**Table 3.2: Usability Scale**

The Evaluators will rate the usability of the system with the scale ranges from 1 to 4 that was elaborated at table 3.1. After grading the system, the proponents will calculate the usability of the system using the weighted mean to get the average score of your system in the usability of the application.

35

**Chapter IV**

**RESULTS AND FINDINGS**

The outcome of the project after the whole development process is discussed in this chapter.

**Project Description**

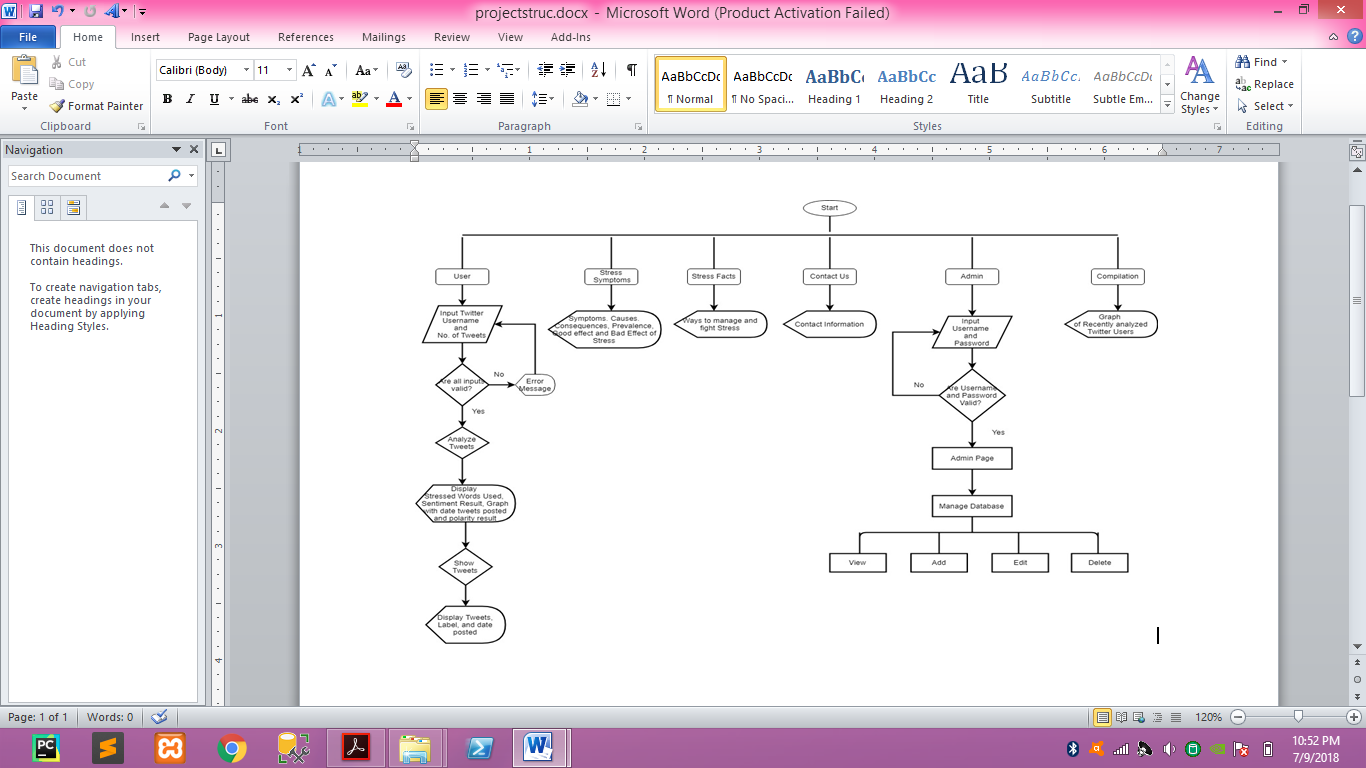
In this project, the proponents have developed a stress recognition system to automatically recognize individual’s psychological stress via social media with the use of a micro-blog called Twitter which is one of the most popular social media that can be publicly accessed that allows individuals to express their opinions, feelings, and thoughts on a variety of topics in the form of short text messages. People can post text with no more than 180 words or have social interactions with others. Employing real online micro-blog data, we investigate the correlations between users’ stress level and their tweeting content by means of Natural Language Processing’s Sentiment Analysis techniques and Information Extraction. The objective of this study is to develop a web based application that will help individuals to be aware of the stress in advance, and manage the stress before it becomes severe and starts causing health problems. The information regarding stress is gathered from many reliable psychological sites and dictionaries. For an additional purpose of this study, we also created a method for automatically rating emotions expressed by Twitter messages. A system developed based on this method could potentially be employed in a large variety of applications, ranging from well-being apps, self-helps, counselors, to

36

community population studies. This project can also be used by healthcare professionals or counseling agencies to monitor and track a patient's emotional states, or to recognize anxiety or systemic stressors of populations (e.g. different student groups on campus). The system can also help commercial agencies to gauge sentiment of buyers or to facilitate targeted product advertisement. In addition, this technology can measure public mood of people in a community, which may help social scientists to understand the quality of life of populations.

**Project Structure**

The part of this project describes how the project looks, its features, dimensions, and specifications.



37

**Figure 4.1 Project Structure**

**Pseudocode**

**Tweets Analysis**

**Input:**

Vocabulary

**Output:**

Classified Tweets

**Process:**

1. **CREATE** SVM problem (Vocabulary)
2. SVM problem = selected Vector and Vocabulary
3. **LET** Node = SVM node
4. Node = SVM Node(Vocabulary count list)
5. **LET** Words = String Split Options
6. **REMOVE** Empty Entries
7. **FOR** ( i =ϴ, i<vocabulary count, i+1)

|  |  |
| --- | --- |
| 8: | **LET** Occurrence count = words count and vocabulary count. |
| 9: | **IF** Occurrence count =ϴ |
| 10: | **THEN** continue |

1. **ADD** Node = new SVM Node
2. **LET** Index = Vocabulary count + 1

38

1. **LET** Value = Occurrence Count
2. **RETURN** Node

**Input:**

Tweets

**Output:**

Analysis Results

**Process:**

1. **LET** problem builder = new text classification
2. **CREATE** problem = Array(word keys, word values)
3. **LET** SVC model = SVC (problem, Kernel Helper, Linear Kernel, 1)
4. **LET** Tweets = (from tweet in twitter context status where status type == USERAnd Screenname = Username)
5. **GET** tweet list count
6. Tweetlist = new list()
7. **FOR** each tweet In Tweetlist

|  |  |
| --- | --- |
| 8: | **GET** tweet |
| 9: | **GET** tweet date |

10: **ADD** tweet

1. **FOR** each tweet in tweelist

9: **LET** Double collective score =ϴ

1. **LET** X = Text Classification Problem Builder
2. **CREATE** node array (word, all words keys)

39

1. **LET** Val = Predict Model (New X)
2. **IF** Val <ϴand New X count >ϴ

14: **THEN** output Stress words

1. **LET** Collective Score = collective Score -1
2. **IF** Val >ϴ

17: **THEN** output Positive words

1. **LET** Collective Score = Collective Score + 1
2. **IF** collective score <ϴ

20: **THEN** Tweet as Stressed Tweet = true

1. **ELSE** Tweet as Stressed Tweet = false

22: **ADD** to tweet list

**Input:**

Tweets

**Output:**

Sentiment Percentage Score

1. **LET** Score = Num (Math.Ceiling(decimal) (Decimal) tweet list
2. **WHERE** x = x and Stressed == true And Count() / (Decimal) count tweet list \*100
3. **IF** Score <= 30
4. **ELSE IF** Score > 30 and Score <=60

6: **THEN** Emotion = Neutral

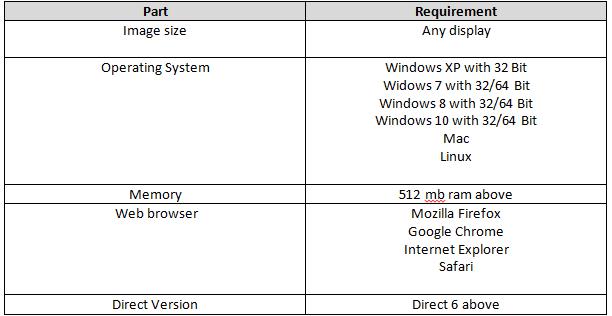
40

1. **ELSE IF** Score > 60

8: **THEN** Emotion = Sign of Stressed

**Technical Specifications**

The proponents recommend using these following technical specifications when using the application for the best performance of the system.



**Table 4.1 Web Application Requirements**

Browser compatibility must be observed to run smoothly the system with all major browsers, such as Mozilla Firefox, Google Chrome, Internet Explorer, Safari, and the like. Stability and expansion of the system should also run with most of the variety of different operating systems and configurations.

41

**Project Capabilities and Limitation**

Psychological stress detection remains a large problem at the present stage. Detecting and managing stress before it turns into severe problems is of significant importance. Recent decades, many efforts have been devoted to stress detection by researchers from diverse areas. They have developed many methods to measure psychological stress, including psychological questionnaire based interviews, self-reported surveys, Forum membership and physiological signal based measuring stress. However, these methods have their limitations in many aspects. Psychological questionnaires and surveys often contain a range of questions designed by psychologists. People are usually unwilling to do these questionnaires unless they have to and this process is time-consuming and costly. Physiological methods usually require professional devices to measure users’ physiological and biochemical properties and need specialists to analyze the acquired data. Thus, it is very important and useful to find a way to detect user’s stress state reliably, automatically and non-invasively.

With the help of natural language processing’s technology and support vector machine algorithm, this project is capable to use social media data to automatically analyze individuals’ stress rate via social media. This project can analyze a maximum of 20 tweets. By clicking the analyze button at the homepage the users can analyze their twitter accounts and the system will automatically direct them to the page of the application where it has the percentage results of all the tweets, the graph with specific dates of their tweets, the negative words they used, and by clicking the “show tweets” button, the

42

users will be able to see if their tweets are positive and negative. This application also provides a site link that has some helpful information for stress people, such as causes, symptoms, prevalence, and treatment of stress. This application also has a page with hotlines that gives free counseling for people who are feeling depressed, stressed and lonely that needs support and a free, anonymous, confidential online text chat with compassionate trained listeners, counselors, and online therapists for users to develop new skills to solve their problems. This application analyzes not just ordinary text but also retweets, mentions, and captions in Twitter. This project also has limitation because there are limitations to what NLP algorithms can handle today. For instance, the tweeted phrase \You're killing it!" may either mean \You're doing great!" or \You're a terrible gardener!" No automated sentiment analysis that currently exists can handle this level of nuance. Furthermore, certain expressions (\ima") or abbreviations (\#lmfao") fool the program, especially when people have maximum 140 characters to express their opinions, or when they use slang, profanity, misspellings and neologisms. This project is also limited to English words only.

**Project Evaluation**

**Feature Selection**

The first objective of the study is to identify the features used to classify the tweets. The factors in analyzing the stress rate of tweets is the sentiment score of every tweets. In order to measure the sentiment score of a tweet, we represent

43

each tweet into a vector of features. We need to capture features that describe the emotion expressed by each tweet. Feature selection plays an important role in the effectiveness of the classification process. For this study, we explore the usage of different features. We use single words, also known as unigrams as the baseline features for comparison. Other features explored included the presence of emoticons, punctuations, capitalization and negations, as elaborated below:



**Figure 4.2: Sample tweet with the used features**

**Unigram features**

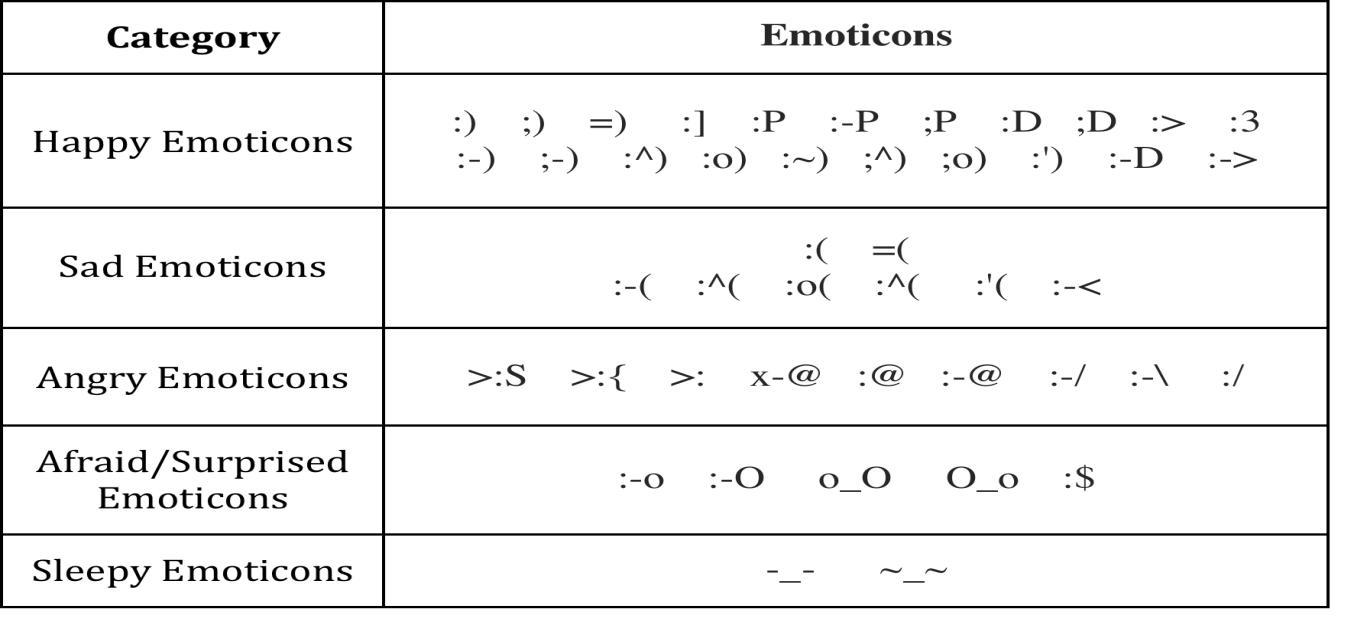
Unigrams or Single word features have been widely used to capture the sentiment of a tweet (Let (f1, f2, ..., fm) be our predefined set of unigrams that can appear in a tweet. Each feature fi in this vector is a word from the dictionary of words in our dataset. Text messages can be classified into emotion categories based on the presence of affect words like ”annoyed”, and ”happy”.

44

Therefore, the problem of high dimensional feature vector can be solved by identifying an appropriate emotion lexicon. We effectively design a domain-specific dictionary by using the lexicon of emotions, instead of all the words in our input dataset.

**Emoticon features**

Other than unigrams, emoticons are likely to be useful features for emotion classification in text messages since they are textual portrayals of a writer’s emotion in the form of icons.



**Table 4.2: Emoticons**

These features tend to be widely used in sentiment analysis. For example, ”:)” and ”:-)” both express happy emotion. The full list of emoticons that we used can be found in Table 4.5

45

**Negation features**

The need negation detection in sentiment analysis can be illustrated by the difference in the meaning of the phrases, "This is good" vs. "This is not good" However, the negations occurring in natural language are seldom so simple. Handling the negation consists of two tasks – Detection of explicit negation cues and the scope of negation of these words.

**Hashtag features**

Twitter message features such as hash-tags and emoticons are likely to be useful features for sentiment and emotion classification. The usage of hashtags in tweets is very common, and Twitter dataset contains millions of different user-defined hash-tags. A study of a sample of 0.6 million tweets by ( Wang et al., 2015) showed that 14.6% of tweets in their sample had at least one hashtag.

**Internet Slang Acronyms**

Slang is a type of language of non-standard words and phrases (Wikipedia, 2014), such as GR8, SMH, CHALE and XOXO. The primary motivation behind the using of Slang words is its usefulness, because usually easy for other to interpret and save a lot of time. Large number of Slangs with positive or negative sentiments are used in chat, Twitter and Facebook messages (Asghar MZ et al., 2014). It has become very important to detect, translate and identify Slang’s polarity for determining the sematic orientation (SO) of the entire review.

46

|  |  |  |  |
| --- | --- | --- | --- |
|  | Slang | Meaning | Orientation |
|  |  |  |  |
|  | **Alr** | Alright | Positive |
|  |  |  |  |
|  | **Damn** | Condemn/Disbelief | Negative |
|  |  |  |  |
|  | **Haha** | Laughing | Positive |
|  |  |  |  |
|  | **Smh** | So much hate | Negative |
|  |  |  |  |
|  | **Tbh** | To be honest | Positive |
|  |  |  |  |
|  | **Xoxo** | Hugs and Kisses | Positive |
|  |  |  |  |
|  | **ROFL** | Rolling on the floor laughing | Positive |
|  |  |  |  |

**Table 4.3: Internet Acronyms**

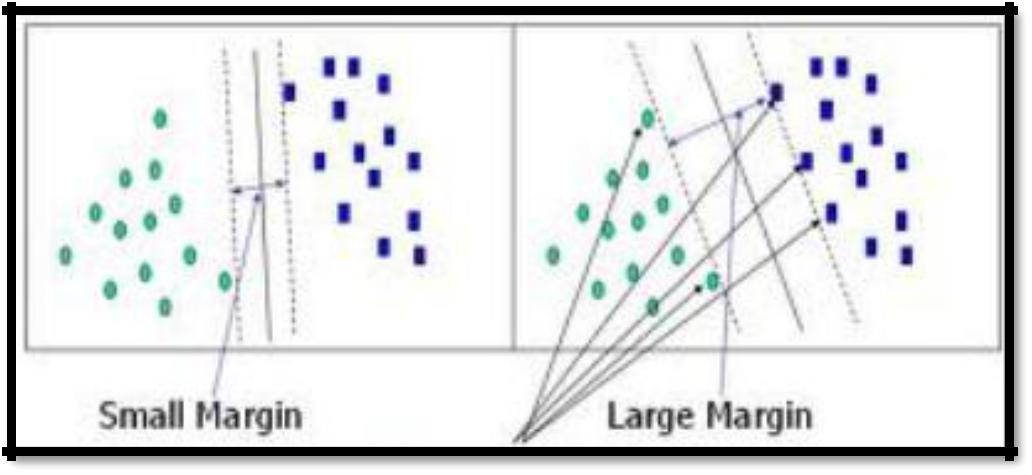
To count the number of polarity words, i.e., Positive and Negative words, we first need to define which words are positive and which words are negative. For this we used AFINN-165. AFINN is a list of English words rated for valence with an integer between -5 (negative) and +5 (positive). The words are manually labeled by Finn Arup Nielsen. The AFINN-165 contains a list of 3382 words and phrases. We process AFINN-165 to obtain 4 features. vNegativeTerms (score -5, -4), NegativeTerms (score -3, -2, -1), PositiveTerms (score 1, 2, 3), and vPositiveTerms (score 4, 5). To count polarity words, documents/tweets are first tokenized. Tokenization means to break a sentence into words. In our implementation, tokens are split by space, punctuation and numbers. For example, the sentence \Been awake since 6am for no reason" contains 8 tokens

47

(see Table 4.4). Once we have tokens, we can simply count the polarity words among these tokens and store them as feature counts.

**Support Vector Machine Algorithm**

To prove the **third objective** that is to implement algorithm and techniques that will analyze signs of stress from twitter data. The proponents used Linear Support Vector Machine Algorithm for the text classification of every tweet.



**Figure 4.3 : Support Vector Machine**

For the text classification we used the supervised learning algorithm called Support Vector Machine. We use the SVMlight software with a linear kernel. Our input data are 3 sets of vectors of size m. Each entry in the vector corresponds to the presence a feature. For example, with a unigram feature extractor, each feature is a single word found in a tweet. If the feature is present, the value is 1, but if the feature is absent, then the value is 0. We use feature presence, as opposed to a count, so that we do not have to scale the input data, which speeds up overall processing.

48

**Confusion Matrix Result**

This part proves the proponents last objective that is to evaluate the

accuracy of the developed application in terms of precision, f-measure,

recall.

**Result**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **@patrickstar15** |  |  |  |  | **Actual Class** | |  |  |  |
|  | **02** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Positive** | | **Negative** | **Neutral** | |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | **Positive** | 5 |  | 0 | 0 |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **Predicted** |  |  | **Negative** | 0 |  | 8 | 2 |  |  |
|  |  |  |  |  |  |  |  |  |
|  | **Class** |  |  |  |  |  |  |  |  |  |
|  |  |  | **Neutral** | 3 |  | 3 | 6 |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |

**Table 4.4 Data for Confusion Matrix**

**Precision:**

**0.31 = 31%**

**Recall:**

49

**F-measure**:

**Accuracy:**

**0.62 = 62%**

**0.41 = 41%**

**0.76= 76%**

50

**Usability**

The respondents of the usability evaluation are the Registered Psychometricians who has a great knowledge about our study which is related to Psychology. The proponents used an evaluation form to get the usability of the system. The formula used to get the usability result is:

**Usability** =

**Result**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Evaluator** |  |  |  |  | **Usability** | |  | **Grade** |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Ms. Glaiza J. Aquino,** |  |  |  |  |  |  |  |  |  |  |
|  | **Registered** |  |  |  |  |  |  |  |  |  |  |
|  | **Psychometrician** |  |  |  |  |  |  |  | **Good** |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

**3.20**

**Table 4.3: Evaluator 1 Result**

The usability of the system is evaluated by Ms. Glaiza J. Aquino, a Registered Psychometrician. The weighted mean of the evaluation is 3.20 which have an equivalent grade of good.

51

|  |  |  |  |
| --- | --- | --- | --- |
| **Evaluator** |  | **Usability** | **Grade** |
|  |  |  |  |

**Ms. Sherica Theriz J.**

**Pasion**



|  |  |
| --- | --- |
| **Registered** | **Excellent** |

**Psychometrician**



**3.60**

The usability of the system is evaluated by Ms. Sherica Theriz J. Pasion, a Registered Psychometrician. The weighted mean of the evaluation is 3.60 which have an equivalent grade of excellent. The formula to get the overall mean of the system is:

**Total Mean =**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Items** | **Mean** | **Grade** |
|  |  |  |  |
|  |  |  |  |
|  | The web application is easy to use. | 4 | Excellent |
|  |  |  |  |
|  | Analysis results are accurate. | 2.5 | Fair |
|  |  |  |  |
|  | Information about stress is correct. | 3 | Fair |
|  |  |  |  |
|  |  |  | Excellent |
|  | This system will be able to help users to be aware of the |  |  |
|  | stress in advance, and manage the stress before it | 4 |  |
|  | becomes severe and starts causing health problems. |  |  |
|  | . |  |  |
|  | The text and information are understandable/easy to read. | 4 | Excellent |
|  |  |  |  |
|  | Gives clear, concise, and consistent labels for navigation. | 3 | Good |
|  |  |  |  |

52

|  |  |  |  |
| --- | --- | --- | --- |
|  | The visual look of the application meets the stylistic | 3.5 | Excellent |
|  | expectations of the target audience. |  |  |
|  |  |  |  |
|  | The text and paragraphs concise and does not contain | 3.5 | Excellent |
|  | needless words. |  |  |
|  |  |  |  |
|  | The alignment establish an appropriate hierarchy of | 3 | Good |
|  | information. |  |  |
|  |  |  |  |
|  | The content is consistent throughout the application. | 4 | Excellent |
|  |  |  |  |
|  | **Total Mean and Overall Equivalent Grade** | **3.79** | **Excellent** |
|  |  |  |  |

**Table 4.5: Overall Usability Result**

The overall mean is 3.79 which have an equivalent grade of Excellent.

**Chapter V**

**SUMMARY, CONCLUSION AND RECOMMENDATION**

This chapter represents the summary, conclusion and recommendation of this study.

**Summary**

Advances in natural language processing and machine learning are making the prospect of large-scale screening of social media for at-risk individuals a near-future possibility. This paper focuses on the broadest view of digital health with a comprehensive scheme for stress level estimation based on social media information. This study makes it feasible to analyze stress and other emotions on several online environments, specially twitter.

Our technique serves as an extra diagnostic tool for psychologists to help detect stress and monitor the emotions of patients, this developed application will also help every individual especially teenagers who are more active on social

53

media to be aware of the stress in advance, and manage the stress before it becomes severe and starts causing health problems. To develop the system that can analyze stress from Twitter, the proponents used Natural Language Processing’s Sentiment Analysis with the help of a supervised machine learning algorithm called Support Vector Machine Algorithm for text classification. The proponents also used sentiment dictionaries like AFINN and Vader Sentiments with a total of 7000 words including emoticons, slangs, and hashtags for a more accurate analysis result. Analysis Results includes the Stressed words used by the twitter account, Sentiment Result, Percentage Rate, a Graph with the date and a total polarity count of every tweets, It also has a show tweets button that will show the recent tweets, the time and date it was published. The application also has a compilation feature that will compile the recent groups that was recently analyzed which may help social scientists to understand the quality of life of populations and measures public mood of people in a community, The system also provides facts and information about stress that can help every user to be aware of stress before it becomes severe and turns to depression. The application has proved to be useful with a usability grade of excellent that was evaluated by Registered Psychometricians and accuracy result of 76%.

**Conclusion**

Mining and analysis of social media activity in order to understand a variety of public health phenomena has been gaining considerable traction

54

recently among researchers. In this paper, we have demonstrated the potential of using social media data as a reliable tool for analyzing stress. The objective was to offer a platform which is fast, accurate and flexible to identify users and analyze patterns of their tweets in terms sentiment and words they are using. With the used of Support Vector Machine Algorithm, Sentiment Dictionaries, NLP’s sentiment Analysis and Information Extraction, the developed Deepter approach enables us to classify large amounts of short texts with no manual effort. The effectiveness of the approach is verified by the experiments. The current study confirms that Twitter is used by individuals not only to expressed they are stressed but the Twitter user’s today also uses the social media to express suicidality/depression especially teenagers who are more active on Twitter.

**Recommendation**

For our future work we recommend to also classify Tagalog words as Filipinos are always number 1 in terms of time spend on social media (Inquirer.net, 2017) and the number of active Twitter users is projected to reach 10.4 million, up from 5.5 million in 2014 (statista.com, 2017). We also recommend analyzing sarcasm and investigating patterns of major stress behavior over weekly or monthly basis for a higher accuracy. Lastly we recommend analyzing depressed users in Twitter since depression today is very common (Petty, 2017). We also recommend adding image recognition as stress

55

tweets today always contains black and white images. Further, we also suggest using community features provided by Twitter API to map data at population level to study trends such as investigating if a certain region contains more users at-risk of mental illness, or if users of certain age are more prone to stress.

**REFERENCES**

Jeanne Segal, Ph.D., Melinda Smith, M.A., Robert Segal, M.A., and Lawrence Robinson. Last updated: February 2018.

Petty (2017). Why Is Everyone I Know Depressed?. Retrieved Dec 21, 2017 from https://greatist.com/live/why-is-everyone-i-know-depressed

Segal, J., Smith, M., Segal, R., Lawrence, R. (2018, February). Stress Symptoms, Signs, and Causes. Retrieved from URL

Efthymios Kouloumpis, Theresa Wilson, and Johanna Moore, “Twitter sentiment analysis: The good the bad and the omg!,” in International AAAI Conference on Weblogs and Social Media (ICWSM’15). 2015, The AAAI Press.

Bagroy S, Kumaraguru P, De Choudhury M: A social media based index of mental well-being in college campuses. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017

56

Gkotsis G, Oellrich A, Hubbard TJ, Dobson RJ, Liakata M, Velupillai S, Dutta R: The language of mental health problems in social media. Third Computational Linguistics and Clinical Psychology Workshop (NAACL). 2016:63-73.

De Choudhury M, Kiciman E, Dredze M, Coppersmith G, Kumar M: Discovering shifts to suicidal ideation from mental health content in social media. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016:2098- 2110.

Liu, B. (2016). Sentiment analysis and subjectivity. Handbook of natural language processing, 2, 627-666. WHO (2016). Mental health: a state of well-being.

Jashinsky, J., Burton, S. H., Hanson, C. L., West, J., Giraud-Carrier, C., Barnes, M. D., et al. (2015). Tracking suicide risk factors through Twitter in the US. Crisis, 35(1), 51e59.

Diana MacLean, Sonal Gupta, Anna Lembke, Christopher Manning, and Jeffrey Heer. 2015. Forum77: An Analysis of an Online Health Forum Dedicated to Addiction Recovery. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW ’15, pages 1511– 1526, New York, NY, USA. ACM.

57

Perrin, A. (2015, October 8). Social Media Usage: 2005-2015.Retrieved from http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015

Yiping Li, Jing Huang, Hao Wang, Ling Feng ‘Predicting Teens’ Future Stress Level from Microblog’, 2015

Newton, C. (2017, September 26). Twitter just doubled the character limit for tweets to 280

.Retrieved from https://www.theverge.com/2017/9/26/16363912/twitter-character-limit-increase-280-test

Segal, J., Smith, M., Segal, R., Lawrence, R. (2018, February). Stress Symptoms, Signs, and Causes. Retrieved from URL

Huijie Lin, Jia Jia, Quan Guo, Yuanyuan Xue, Jie Huang, Lianhong Cai, and Ling Feng. Psychological stress detection from cross-media microblog data using deep sparse neural network. In Proceedings of International Conference on Multimedia and Expo, pages 1–6. IEEE, 2014.

Yuanyuan Xue, Qi Li, Ling Feng, Gari D Clifford, and David A Clifton. Towards a micro-blog platform

for sensing and easing adolescent psychological pressures. In Proceedings of the ACM conference on Pervasive and ubiquitous computing adjunct publication, pages 215– 218. ACM, 2013.

58

Selye and others, 2016 Hans Selye et al. A syndrome produced by diverse nocuous agents. Nature, 138(3479):32, 2016.

Lin et al., 2014b Huijie Lin, Jia Jia, Quan Guo, Yuanyuan Xue, Qi Li, Jie Huang, Lianhong Cai, and Ling Feng. Userlevel psychological stress detection from social media using deep neural network. In Proceedings of ACM International Conference on Multimedia, pages 507–516. ACM, 2014.

Glen Coppersmith, Craig Harman, and Mark Dredze. Measuring post traumatic stress disorder in twitter. In Proceedings of the

International Conference on Weblogs and Social Media, pages 579–582, 2014.

Rui Fan, Jichang Zhao, Yan Chen, and Ke Xu. Anger is more influential than joy:

Sentiment correlation in weibo. PLoS ONE,

2014.

H. Lin, J. Jia, Q. Guo, Y. Xue, J. Huang, L. Cai, and L. Feng. Psychological stress detection from cross-media microblog data using deep sparse neural network. In proceedings of IEEE International Conference on Multimedia & Expo, 2014.

59

Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. Sentiment strength detection in short informal text. Journal of the American Society for Information Science and Technology, 61(12):2544–2558, 2015.

Yuan Zhang, Jie Tang, Jimeng Sun, Yiran Chen, and Jinghai Rao. Moodcast:

Emotion prediction via dynamic continuous factor

graph model. 2013 IEEE 13th International Conference on Data Mining, pages 1193–1198, 2015.

Jichang Zhao, Li Dong, Junjie Wu, and Ke Xu. Moodlens: an emoticon-based sentiment analysis system for chinese tweets. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 1528–1531, 2013.

Golnoosh Farnadi, Geetha Sitaraman, Shanu Sushmita, Fabio Celli, Michal Kosinski, David Stillwell, Sergio Davalos, Marie Francine Moens, and Martine De Cock. Computational personality recognition in social media. User Modeling and User- Adapted Interaction, pages 1–34, 2016.

Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. Predicting personality from twitter. In Passat/socialcom 2011, Privacy, Security, Risk and Trust, pages 149–156, 2011.

60

Ben Verhoeven, Walter Daelemans, and Barbara Plank. Twisty: A multilingual twitter stylometry corpus for gender and personality profiling. In Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC, pages 1632–1637, 2016.

Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Pentland. Daily stress recognition from mobile phone

data, weather conditions and individual traits. In ACM International Conference on Multimedia, pages 477–486, 2014.

Jichang Zhao, Li Dong, Junjie Wu, and Ke Xu. Moodlens: an emoticon-based sentiment analysis system for chinese tweets. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 1528–1531, 2012.

Rui Fan, Jichang Zhao, Yan Chen, and Ke Xu. Anger is more influential than joy:

Sentiment correlation in weibo. PLoS ONE, 2014.

Lu, H., Frauendorfer, D., Rabbi, M., Mast, M. S., Chittaranjan, G. T., Campbell, A. T., ... & Choudhury, T. (2012, September). StressSense: Detecting stress in unconstrained acoustic environments using smartphones. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (pp. 351-360). ACM.

61

Paredes, P., Sun, D., & Canny, J. (2013, May). Sensor-less sensing for affective computing and stress management technology. In Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on (pp. 459-463). IEEE.

Acar Tamersoy, Munmun De Choudhury, and Duen Horng Chau. 2015. Characterizing Smoking and Drinking Abstinence from Social Media. In Proceedings of the 26th ACM Conference on Hypertext & Social Media, pages 139–148, Guzelyurt, TRNC, Cyprus, September 1-4th 2015. ACM.

Stevie Chancellor, Tanushree Mitra, and Munmun De Choudhury. 2016. Recovery Amid Pro-Anorexia. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI ’16, pages 2111–2123, New York, New York, USA. ACM Press.

Xie, Y., Chen, Z., Cheng, Y., Zhang, K., Agrawal, A., Liao, W.-K., and Choudhary, A. (2013). Detecting and tracking disease outbreaks by mining social media data. Dimensions.

Lin, H., Jia, J., Guo, Q., Xue, Y., Li, Q., Huang, J., Cai, L., Feng, L.: User-level psychological stress detection from social media using deep neural network. In: Proceedings of the 22nd ACM International Conference on Multimedia. pp. 507– 516.

62

Philip Resnik, William Armstrong, Leonardo Claudino, and Thang Nguyen. 2015a. The University of Maryland CLPsych 2015 Shared Task System. In CLPsych 2015 Shared Task System. c,pages 54–60.

Acar Tamersoy, Munmun De Choudhury, and Duen Horng Chau. 2015. Characterizing Smoking and Drinking Abstinence from Social Media. In Proceedings of the 26th ACM Conference on Hypertext & Social Media, pages 139–148, Guzelyurt, TRNC, Cyprus, September 1-4th 2015. ACM.

Olutobi Owoputi, Chris Dyer, Kevin Gimpel, and Nathan Schneider. 2012. Part-of-Speech Tagging for Twitter: Word Clusters and Other Advances. Technical report, Carnegie Mellon University.

Mike Conway and Daniel O’Connor. 2016. Social Media, Big Data, and Mental Health: Current Advances and Ethical Implications. Current Opinion in Psychology, 9:77–82. CMHA (2016). Facts about mental illness.

Daniel Preotiuc-Pietro, Johannes Eichstaedt, Gregory Park, Maarten Sap, Laura Smith, Victoria Tobolsky, H. Andrew Schwartz, and Ungar Lyle. 2015. The Role of Personality, Age, and Gender in Tweeting about Mental Illness. In e 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality,, pages 21–23. Association for Computational Linguistics.

63

Farnadi, G., Zoghbi, S., Moens, M.-F., and De Cock, M. (2013). Recognising personality traits using facebook status updates. In Proceedings of the workshop on computational personality recognition (WCPR13) at the 7th international AAAI conference on weblogs and social media (ICWSM13). AAAI.

Farzindar, A. and Inkpen, D. (2015). Natural language processing for social media. Synthesis Lectures on Human Language Technologies, 8(2):1{166. Gerber, M. S. (2014). Predicting crime using twitter and kernel density estimation. Decision Support Systems, 61:115125.

Gottschlich, J., Heimbach, I., Hinz, O., et al. (2013). The value of users' facebook pro\_le data-generating product recommendations for online social shopping sites. In ECIS, page 117.

Han, J., Pei, J., and Kamber, M. (2011). Data mining: concepts and tech-

niques. Elsevier. He, H. and Garcia, E. A. (2009). Learning from imbalanced data. IEEE Transactions on knowledge and data engineering.

Moreno, M. A., Jelenchick, L. A., Egan, K. G., Cox, E.,Young, H., Gannon, K. E., and Becker, T. Feeling bad on Facebook: Depression disclosures by college students on a social networking site. Depression and Anxiety 28, 6 (June 2011), 447–455.

64

Park, M., Cha, C., and Cha, M. Depressive moods of users portrayed in twitter. In Proceedings of the ACM SIGKDD Workshop on Healthcare Informatics (HI-KDD’12) (Aug. 2012), 1–8.

Park, S., Lee, S. W., Kwak, J., Cha, M., and Jeong, B. Activities on Facebook reveal the depressive state of users. Journal of Medical Internet Research 15, 10 (Oct. 2013), e217.

Tsugawa, S., Mogi, Y., Kikuchi, Y., Kishino, F., Fujita, K., Itoh, Y., and Ohsaki, H. On estimating depressive tendency of twitter users from their tweet data. In Proceedings of the 2nd International Workshop on Ambient Information Technologies (AMBIT’12) (Mar. 2013), 29–32.

De Choudhury, M., Gamon, M., Counts, S., and Horvitz, E. Predicting depression via social media. In Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM’13) (July 2013), 128–137.

De Choudhury, M., Counts, S., and Horvitz, E. Predicting postpartum changes in emotion and behavior via social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI’13) (Apr. 2013), 3267–3276

65

De Choudhury, M., Counts, S., and Horvitz, E. Social media as a measurement tool of depression in populations. In Proceedings of the 5th Annual ACM Web Science Conference (WebSci’13) (May 2013), 47–56.

Culotta, A. Estimating county health statistics with twitter. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI’14) (Apr. 2014), 1335–1344.

Paul, M. J., and Dredze, M. You are what you tweet: Analyzing twitter for public health. In Proceedings of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM’11) (July 2011), 265–272.

Tejwani, R. (2014). Sentiment Analysis: A Survey. arXiv preprint arXiv:1405.2584. Liu, B. (2010). Sentiment analysis and subjectivity. Handbook of natural language processing, 2, 627-666.

J. Bollen, A. Pepe, and H. Mao, “Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena,” CoRR, vol. abs/0911.1583, 2009.

A. Pepe and J. Bollen, “Between conjecture and memento: shaping a collective emotional perception of the future,” CoRR, vol. abs/0801.3864, 2008.

66

A. Pak and P. Paroubek, “Twitter as a corpus for sentiment analysis and opinion mining,” in Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC’10), N. C. C. Chair), K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis,

M. Rosner, and D. Tapias, Eds. Valletta, Malta: European Language Resources Association (ELRA), may 2010.

Y. Lu, M. Castellanos, U. Dayal, and C. Zhai, “Automatic construction of a context-aware sentiment lexicon: an optimization approach,” in Proceedings of the 20th international conference on World wide web, ser. WWW ’11. New York, NY, USA: ACM, 2011, pp. 347–356.

De Choudhury, M., Gamon, M., Counts, S., Horvitz, E.: Predicting Depression via Social Media. In: Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM). p. 2 (2013).

Daniel Preotiuc-Pietro, Maarten Sap, H. Andrew Schwartz, and Lyle Ungar. 2015. Mental Illness Detection at the World Well-Being Project for the CLPsych 2015 Shared Task. In Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. pages 40–45.

67

Lin, H., Jia, J., Guo, Q., Xue, Y., Li, Q., Huang, J., Cai, L., Feng, L.: User-level psychological stress detection from social media using deep neural network. In: Proceedings of the 22nd ACM International Conference on Multimedia. pp. 507– 516. ACM (2014).

B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up? sentiment classification using machine learning techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP ’02, 2007.

Stewart, A. E. (2014). Vitamin D deficiency, depression linked in UGA-led international study. Retrieved from <http://news.uga.edu/releases/article/vitamin-d-deficiency-depression/>website.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. Journal of language and social psychology, 29(1), 24-54.

Tejwani, R. (2014). Sentiment Analysis: A Survey. arXiv preprint arXiv:1405.2584. Turney, P. D. (2002). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. Paper presented at the Proceedings of the 40th annual meeting on association for computational linguistics.

68

Twitter, I. (2015). Developers Documentation. Retrieved Aug 10, 2015, from [https://dev.twitter.com/overview/documentation.](https://dev.twitter.com/overview/documentation)

Vikram Patel, F. M. S., and Shekhar Saxena, M.D. (2014). Transforming Lives, Enhancing Communities — Innovations in Global Mental Health. The new England Journal of Medicine.

Wang, X., Zhang, C., Ji, Y., Sun, L., Wu, L., & Bao, Z. (2013). A depression detection model based on sentiment analysis in micro-blog social network Trends and Applications in Knowledge Discovery and Data Mining (pp. 201-213): Springer.

69