#ActuallyDepressed: Characterization of Depressed Tumblr Users' Online Behavior from Rules Generation Machine Learning Technique

Czarina Rae C. Cahutay
University of the Philippines Cebu
Gorordo Avenue, Lahug, Cebu City
czarinacahutay@gmail.com

Aileen Joan O. Vicente
University of the Philippines Cebu
Gorordo Avenue, Lahug, Cebu City
aovicente@up.edu.ph

Abstract

The ubiquitous data provided by social networking sites paved the way for researchers to understand netizens behavior with psychological ailments such as depression. However, most of these researches are aimed at classifying users with depression using blackbox algorithms such as SVM. This does not allow data exploration or further understanding the characteristics of depressed individuals. This research aims to characterize depressed Tumblr users online behavior from rules generation . Characterization is done by collecting affective, social, cognitive and linguistic style markers collected from the respondents posts. Rules are then generated from these features using CN2 algorithm — a rules generating machine learning technique. The rules are analyzed and are compared with observations in prior literature on depressive behavior. We observed that depressed respondents in Tumblr have more photo content in posts rather than just pure textual posts, posts are more negative, and there is an evident use of self-referencing words. These characteristics are also evident in offline behavior of depressed individuals based on prior literature

Keywords: Depression, Characterization, Rules Generation, CN2 Algorithm

1. Introduction

According to the World Health Organization (WHO) (2001), depression is a serious mental issue and may be the second leading cause of disease by 2020. Amy Courtney (2014) argues that blogging reduces the symptoms brought about depression. Blogging allows the public to access and comment on such work, allowing additional psychological benefits.

Microblogging is an easier way of blogging that allows users to create short contents shared with an audience in real-time. This creates an easier and more convenient way of sharing content and information through the web (Nations, 2015). Contents vary from text, to visual, audio, audiovisual and even the use of links to redirect to other websites. This study aims to characterize the online behavior of depressed individuals using machine learning technique. Further, this research is guided by the following questions:

RQ1. What are the characteristics of depressed individuals on Tumblr in terms of their affective, behavioral, cognitive, and linguistic style attributes?

RQ2. How do these characteristics compare to offline behavior of depressed individuals in literature? Are they any different?

This study focuses on Tumblr because there have not been a lot of studies focusing on adolescents social media postings. Other studies on depression online are focused on an adult age group. Moreover, we find that because Tumblr allows anonymity of users there might be a more genuine response with regards to their posting and the results that we gather since they are not bound to their true identity in person similar to the study of Warner et al (2016).

The research covers 17 features of depressed individuals that are explored in other studies as classification problems. With these 17 features, 13 are coming from the posts, while the other 4 are basic information about the user (age, civil status, highest education attained and gender).

Despite a number of studies correlating with depression, we extracted features from 4 different processes or style attributes. These are: (1) affective, (2) social, (3) cognitive, and (4) linguistic. In other studies (Reece, 2016; De Choudhury, 2014; Moreno, 2011), only cognitive and linguistic patterns have been extracted. By extending the depression studies to include its social and affective process and allowing a more varied set of attributes, this research aims to create a wholistic characterization of such users.

2. Background and Related Work

2.1. Depression Studies on Social Media

A prominent work by De Choudhury et al (2013) reveals how depression can be predicted in twitter based on linguistic patterns. The researchers crowdsourced labels as ground truth data for Major Depressive Disorder (MDD). Using Amazon's Mechanical Turk interface, they successfully design human intelligence task for crowdworkers to take standardized clinical depression survey.

The researchers used a CESD questionnaire as a primary tool for determining depression levels which is a self-report scale to measure depressive symptoms. The range of the score of this scale varies from 0 - 60.

Another study by Hu et al (2015) explored on Predicting Depression of Social Media User on Different Observation Windows. The study is done to predict a user's depression based on Weibo data, thus building a regression model to predict the CES-D score of any user. The study featured around 900 features but deemed only 200 to be used. The goal of the study by Hu et al (2015) only considered the accuracy and performance of the regression models that the researchers have developed without taking focus on the specific behaviors of such depressed individuals.

2.2. Characterization Methods in Social Media Studies

There are a number of researches that aim to understand or characterize various types of behavior and phenomena on social media. Most of these characterization methods involve deriving patterns from statistical data such as means and p-values.

To describe their posting patterns, they create Cumulative Distribution Functions (CDF) of different variables. They focus on the variables such as popularity, reciprocity, retweet_ratio, url_ratio, mention_ratio, hashtag_ratio. Moreso, they create a heatmap that shows the tweet frequencies for different days and hours on a four-month period. They answer research questions they have identified in their study from their characterization.

A notable characterization study was made by De Choudhury et al (2015) on characterizing anorexia on Tumblr. By using the Tumblr API, they collect 55,334 public language posts from 18,923 unique blogs. They manually examined these posts and obtained a list of 28 tags relating to eating disorder and anorexia.

They characterize Tumblr through affective, social, cognitive and linguistic style processes to determine the features used. By using statistical methods such as p-scores and z-scores, they were able to translate qualitative data into quantitative data.

The features of the depression studies of De Choudhury et al (2013) and Hu et al (2015) focus on blackbox algorithms such as SVM where the understanding of these features are not explored in detail. These aforementioned studies have focused on predicting and not understanding the behavior of such users.

3. Methodology

This research is developed through understanding the behavior of such users through a rule generation machine learning technique using a Sequential Covering Algorithm. The rules generated are analyzed further to get characteristics of the online behavior of depressed individuals.

3.1. Data Collection

Tags such as #depressed, #actuallydepressed, #depressing, #depression, and #actuallyborderline are used to initially gather information about depressive posts. The tags are used to retrieve other tags associated with depression. Tags that do not necessarily characterize or which may not automatically correlate with depression such as #sad, #family, #words are removed from the set of tags so that the tags only pertaining to depression or its types are used. We only include the tags that are associated with depression or its symptoms such as feeling suicidal.

The initial set of usernames are collected through crawling the tags and posts during the initial data collection. An invitation is sent to these pre-processed users to answer the survey which contains the Center for Epidemiologic Studies Depression (CESD) Scale (http://cesd-r.com/) and Primary Health Questionnaire - 2 (PHQ-2) questionnaires. Only processed users who post English content that has posts with tags correlated to depression are used in this study.

The information collected from the users are the posts coming from the first 20 pages of their accounts. The number of posts ranges from 100 to 228 posts per user depending on the individual layouts done by the user on their webpages. Tumblr allows users to customize their profiles according to their liking thus there is no specific number of posts per page.

A set of 20 tags collected from tag-crawling are used to gather information about depressive posts. 447 users are contacted through the instant messaging feature on Tumblr. These users have been initially screened to meet the corpus requirements that they: (1) only post English-content posts, and (2) have posted media in their microblogs with a tag correlating with depression. Of the 447 messaged users, only 130 users responded. It is expected that a larger number of responses will be categorized as depressed as the data collection is aimed at collecting posts from depressed individuals.

The survey responses gathered from 130 respondents are analyzed to be able to ascertain that the respondents are really depressed. From the 130 respondents, 70 are actually depressed, 19 are not depressed , and 41 are invalid.

3.2. Feature Extraction

Affective, cognitive, social and linguistic features are retrieved from the collected data of these users.

Affective processes are feelings, responses (usually quantified as positive or negative), emotion-laden behavior and beliefs (StateUniversity.com, 2016). For affective processes, the positive, negative and neutral reaction from data collected from these users is measured through subjectivity and positivity. Using the NTLK library from python, we attach subjectivity and positivity scores to the dataset. Positivity is a range from -1 to 1 determining the emotion generated from the text. A lower positivity score would mean that the post shows more negative emotion a post has based on the NTLK. Subjectivity is based on the emotive expression being detected from the posts. It is a range from 0 to 1, where 0 determines that the post is more objective.

For cognitive processes, we measure the user posting and behavior. There are six types of measures that are used: (1) self-referencing (I, me, my, etc.) which is also one of the many signs and symptoms of depression (Segrin, 2000), (2) average number of articles (a, an, the) a feature also used in the study of Hu et al (2015), (3) big words defined as greater than 6 letters, also used in the study of Hu et al (2015). We also consider the words that are (4) Social and personal concerns pertaining to, family, friends, work and home.

Social Processes are collected through the following measures: (1) gender of the individual, (2) level of education, (3) civil status, as these three components are linked to determining if the person is more prone to depression based on prior research (Ross & Mirowsky, 1989), (4) Number of photo posts (Segrin, 2000) is also collected as this determines whether these individuals prefer posting content with photos, (5) the average difference between each post which determines how much the user interacts with the set of followers that he has, and determines how often he shares content.

Linguistic Features are extracted through (1) verbal fluency as the number of words in a post, (2) number of sentences in a post, (3) number of unique tags in a post which would determine how specific the user is using the tags to determine a specific post.

3.3. Characterization Using Rule Extraction

This study will explore rule extraction through a Sequential Covering Algorithm (CN2) for rule generation. A free open-source software, Orange (https://orange.biolab.si/) developed by the University of Ljubljana, is used to implement the algorithm. The algorithm is provided with nominal and numeric features and a target variable of 0 or 1 that indicates whether or not the user is depressed.

The CN2 Algorithm was developed by Clark and Niblett in 1989 as an improved version of the ID3 and AQ algorithm that are used for rules generation and tree generating algorithms. CN2 algorithm uses entropy to determine the best complex found. A complex is a condition that when satisfied, will minimize the number of examples that the algorithm needs to explore to determine the label of the class. Rules are determined by an if-then condition where the complex and labels complete the condition, if <complex> then predict <label> where <complex>. The algorithm determines the new complex by finding the set of examples that the complex covers. The entropy function prefers a large number of examples of a single class with a few examples of another class, resulting that these complexes perform well on the training data.

CN2 rule induction results in an ordered or unordered list of if-then rules, removing the items in the training data that are captured by the consecutive set of rules, finishing off with a sequence of if-then statements that determine the label of the data. The heuristics used by the CN2 algorithm uses entropy to determine the *best complex* found. A complex is a condition that when satisfied, will minimize the number of examples that the algorithm needs to explore to determine the label of the class. Rules are determined by an if-then condition where the complex and labels complete the condition, if *<complex>* then predict *<label>* where *<complex>*. The algorithm determines the new complex by finding the set of examples that the complex covers.

3.4. Rule Validation and Rule Interpretation

Orange (http://orange.biolab.si/) has a built-in Test and Score function that offers a “leave one out” testing method. Although the study is not aiming for classification, the validation is needed to verify if the rules generated by the CN2 Algorithm are quality rules with good performance based from the data. The entire dataset is used in both training and testing due to the limited number of respondents.

The testing validation used by the Orange (http://orange.biolab.si/) software returns recall, precision and F1-score. The three being the most commonly used testing validations for classification problems. This would also help in determining our confidence with the rules that were derived by the CN2 Algorithm to correctly understand the data.

3.5. External Validation

The general patterns on the findings of the data are discussed in detail and and understood in the context of depression. An attempt to compare such retrieved behavior with offline behavior of the patients is also be explored to answer the research question presented prior to the specific objectives.

4. Results and Discussion

4.1. Dataset Description

Based on the collected data from Tumblr users, most of the depressed Tumblr users are at the ages of 16-19. The distribution of these respondents can be found in Figure 4.1.

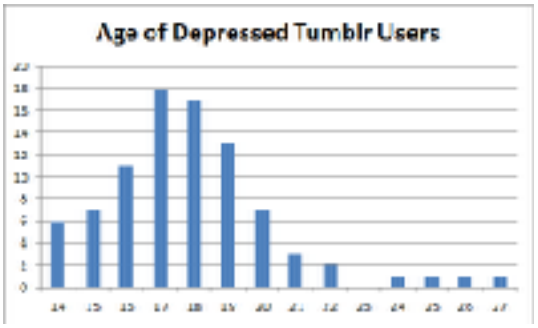

Figure 4.1. Age distribution of Depressed Tumblr Users

The age distribution would suggest that most of these users lie at the high school age group. Most of the depressed user respondents high school level as their highest educational degree attained.

It would also hold that these depressed Tumblr users would have similar civil status. Thus, it shows that only 2 out of the 70 depressed users are married while the rest are single.

The gender distribution of the depressed respondents is found in Figure 4.2. Depressed respondents are predominantly female. It is followed by other genders composed of the following genders: bisexual, agender, genderfluid, agendertransexual, genderflux, non-binary males and females, and cisgender.

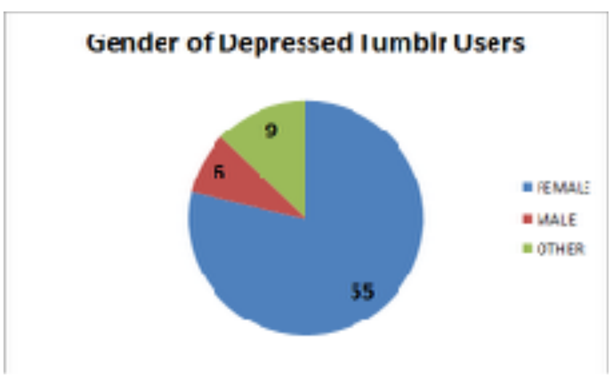

Figure 4.2 Different Genders of Depressed Tumblr Users

General descriptions of the features used in this study are described in the succeeding tables.

Table 4.1 Affective Processes: Polarity and Subjectivity of User Posts

	Polarity	Subjectivity	Polarity Range	Subjectivity Range
Non-Depressed	0.108752	0.300921	-0.00672 to 0.307344	0.06098 to 0.502567
Depressed	0.05698	0.22518	-0.04871 to 0.150953	0.02293 to 0.465565

Table 4.2 Social Processes: Posting Frequency and Photo Use

	Average Time Difference	Average percentage of posts with photo	Average number of photos per photo post
Non-Depressed	1303.013	0.528530	1.011022
Depressed	1165.601	0.625671	1.019832

Table 4.3 Cognitive Processes: Self-referencing, Articles and Themes Used

	Average self-referencing words	Average articles used	Average number the theme was mentioned
Non-Depressed	1.196597	1.977348	0.017669
Depressed	0.782567	1.154454	0.049233

Table 4.4 Linguistic Processes: Form and Context of User Posts

	Average tags per post	Average number of words more than 6 letters	Average number of sentences per post	Average number of words per post	Total number of tags
Non-Depressed	3.44023	11.90891	2.260369	41.22556	137.8421
Depressed	1.95096	7.700895	1.582773	26.37325	86.1

4.2. Respondents

Of the 447 messaged users, only 130 users responded. It is expected that a larger number of responses would be categorized as depressed as the data collection is aimed at collecting posts from depressed individuals.

Based on the analysis of the survey responses of the 130 respondents, 70 are actually depressed, 19 are not depressed and 41 were invalid. Quality responses are achieved by two measures:

a. The results for both PHQ-2 and CESD-R should be the same. That is, if a user scores depressed on the PHQ-2 scales, the user must also score similarly on the CESD-R scale.

b. The control question in between the CESD-R questionnaire should be correctly answered to eliminate the responses from respondents who did not read the questions carefully giving dishonest answers.

Other responses have also been considered invalid even if they passed both quality measures because their Tumblr blogs are private and they requested to keep its privacy.

4.3. Tags Used

Although the tags that have been used to retrieve the pool of respondents were related to depression, a number of non-depressed users (19 out of 130 respondents) have also responded.

Figure 4.4a is a tag cloud containing the top 25 tags that are used by depressed users while Figure 4.4b contains that top 25 tags used by non-depressed users. Each tag is associated with a weight that corresponds to the number of times the tag has been used. The larger the tag appears on the tag cloud, the more frequent it has been used.

Depressed users often incorporate tags in their posts that were used in retrieving the usernames of potential respondents. Most of their tags present depressive content and often employs the word depression.

On the other hand, non-depressed users do not have any of the 20 tags that the study used in order to retrieve the potential respondents. The tags are rarely used and are overpowered by the tags found in Figure 4.4b that they are frequently using. Non-depressed users often use depression-associated tags within their posts encouraging depressed users to seek help or when only a limited number of posts are aimed at depression. Some of these non-depressed users may be only mildly depressed depending on the scores of both PHQ-2 and CESD-R questionnaires.

(a) (b)

Figure 4.4 Tag Clouds Generated from the Top 25 Tags Used by (a) Depressed and (b) Non-Depressed Users

4.4. Rules Generated

Six (6) unordered rules are generated from CN2 Rule Induction given maximum coverage of 4% (3 out of 89) and a statistical significance of 1 (default alpha). As this study focuses on the behavior of depressed Tumblr users, only the rules with a class label of 1 are analyzed. The rules can be found in Table 4.6 which contains two columns that describe the rules: (1) if <complex>, (2) then <class>. The complex contains the condition that is to be satisfied in order for one to receive a label of a class.

Table 4.5 Unordered Rules

	IF <complex>	THEN <class>
1	polarity >= 0.163177578	label = 0
2	time elapsed >= 163.8604768	label = 0
3	total tags >= 13	label = 0
4	has image >= 0.8739	label = 1
5	polarity <= 0.090295132	label = 1
6	self referring words <= 3.225	label = 1

Several different types of features were included in the study in the hope of gathering as much information about depressed online users. Apparently, only three (3) features significantly characterizes depressed online users covering affective, social and cognitive behaviors.

The most prominent characteristic based on the rules generated would be that depressed Tumblr users most likely posts with photo content. Reece and Danforth (2016) support this characteristic in their study where users can be found depressed based on their photo posts. Moreso, this study would support this such that depressed individuals also post more images frequently as compared to non-depressed individuals.

Polarity is found as a feature in both depressed and non-depressed individuals. However, a polarity closer to 1 would entail a more positive post while a polarity score closer to 0 would mean a more negative post. From the rules, depressed individuals tend to post more negative posts as compared to non-depressed individuals. The idea of depressive realism explored by Burton (2012) would reflect a depressive characteristic that depressed individuals often have a negative perspective in life allowing them to be more objective. A study by Gotlib and Joormann (2010) find that depression is characterized by an increase in the elaboration of negative information, difficulties in disengaging from negative which may be reflective in their posts.

Lastly, the use of self referring words can also characterize depressed individuals. Despite the depressed Tumblr users average use of self-referring words being 0.7, the threshold retrieved from the rules is at 3. Because of this, it is enough to say that depressed users highly use self-referencing words. There have been linguistic markers of depression as studied by Hargitay, et. al(2007). These linguistic markers are found the self-narratives of depressed individuals that feature the pronouns 'I' and 'myself'. The linguistic markers can be seen as "reflecting pervasive and enduring psychological processes" (Hargitay et al, 2007). Another study by Rude et al (2004), depressed individuals used the word "I" more than never depressed participants.

4.5. Rule Validation Results

While this study is not intended to classify, we find it imperative to assure the quality of the rules. After a leave-one-out cross validation, the classification performance metrics are generated. The summary is found in Table 4.7

Table 4.6. Evaluation Metrics

F1 Score	Precision	Recall
0.885	0.802	0.986

The F1-Score of 88.5% is a relatively good performance metric that can signify a relatively strong confidence in the rules generated and consequently, its interpretation. The skewness of the data where there are 70 depressed users and only 19 non-depressed users may be accounted for the low precision and recall scores. This is because the data gathering is targeted towards users who have posted depressed content based on the gathered tags.

5. Conclusion

Depression is a pressing issue, and will continue to haunt individuals making it one of the possible leading causes of deaths in adolescents by 2020. The study observed the posting patterns of 89 Tumblr users, 70 of which are depressed according to the PHQ-2 and CESD-R scales. The mean age of the individuals who participated in the study was between 16-19 years old.

Therefore, a depressed Tumblr user is characterized by photo content in posts rather than just pure textual posts, posts are more negative, and the evident use of self-referencing words.

The study concludes that it is possible to use a rule generation machine learning technique in characterizing online behavior.

6. Recommendations

This study only covers the characterization of 70 depressed Tumblr users. Moreover, we have not covered linguistic features that are evident in depressed Tumblr users. Increasing the number of respondents and the number of depressed user posts in the study will allow us to cover other processes and features that are not year clear in this study.

Classification problems that have been using black box algorithm methods such as the work of Hu et al (2015) use 200-900 different features. This study only focuses on 17 features so increasing the number of features would allow more characterization rules to be evident.

References

All about depression: Diagnosis. (2013). Retrieved December 7, 2016, from All About Depression, http://www.allaboutdepression.com/dia_03.html

American Psychological Association. (2016). Center for epidemiological studies depression (CESD). Retrieved December 7, 2016, from American Psychological Association, http://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/depression-scale.aspx

American Psychological Association. (2016). Patient health questionnaire (PHQ-9 & PHQ-2). Retrieved December 09, 2016, from http://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/patient-health.aspx

Beattie, G.S. (2005, November). Social Causes of Depression. Retrieved May 31, 2017, from http://www.personalityresearch.org/papers/beattie.html

Burton, N. (2012, June 5). Depressive Realism. Retrieved May 31, 2017, from https://www.psychologytoday.com/blog/hide-and-seek/201206/depressive-realism

Clark, P., Niblett, T. (1988, October 25). The CN2 induction Algorithm. Retireved May 10, 2017, from https://pdfs.semanticscholar.org/766f/e3586bda3f36cbcce809f5666d2c2b96c98c.pdf

De Choudhury, M., Counts, S., Horvits, E., & Hoff, A. (2014). Characterizing and Predicting Postpartum Depression from Shared Facebook Data.

De Choudhury, M., Gamon, M., Couns, S., & Horvitz, E. (2013). Predicting Depression via Social Media.

Gotlib IH, Kasch KL, Traill S, Joormann J, Arnow BA, Johnson SL. (2010) Coherence and specificity of information-processing biases in depression and social phobia. J Abnorm Psychol. 2004;113(3):386-98.

Gotlib, I. H., & Hammen, C. L. (1992). Psychological aspects of depression: Toward a cognitive-interpersonal integration. New York: Wiley.

Gotlib IH, Joormann J. Cognition and depression: current status and future directions. Annu Rev Clin Psychol. 2010;6:285-312.

Hu, Quan, Ang Li, Fei Heng, Jianpeng Li, and Tingshao Zhu. "Predicting Depression of Social Media User on Different Observation Windows." 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) (2015): n. pag. Web.

Koopman, C., Ismailji, T., Holmes, D., Classen, C., Palesh, O., & Wales, T. (2005, March). The Effects of Expressive Writing on Pain, Depression and Posttraumatic Stress Disorder Symptoms in Survivors of Intimate Partner Violence.

Lowe, N. (n.d.).Social Support Networks and Their Role in Depression. Retrieved May 31, 2017, from http://www.personalityresearch.org/papers/beattie.html

Ramirez-Esparza, N., Chung, C.K., Kacewicz, E., & Pennebaker, J.W. (2008). The psychology of word use in depression forums in English and in Spanish: Testing two text analytic approaches. Association for the Advancement of Artificial Intelligence, 102-108.

Reece, A., Reagan, A., LM Lix, K., Dodds, P. S., Danforth, C., & Langer, E. (2016, August). Forecasting the onset and course of mental illness with Twitter data.

Smith, M., Segal, R., & Segal, J. (2016, November). How to recognize the symptoms and get effective help. Retrieved November 15, 2016, from http://www.helpguide.org/articles/depression/depression-signs-and-symptoms.htm

StateUniversity.com. (2016). Individual differences - Affective and Conative processes. Retrieved December 7, 2016, from http://education.stateuniversity.com/pages/2081/Individual-Differences-AFFECTIVE-CONATIVE-PROCESSES.html

Thapar, A. K., Collishaw, S., & Pine, D. S. (2012). Depression in adolescence. , 379(9820), . Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3488279/

World Health Organization. (2001). The World Health Report 2001: Mental Health: New Understanding, New Hope. Retrieved from http://www.who.int/whr/2001/en/whr01_en.pdf