

Examining the Language of Eating Disorders in Social Media: Indications of Anorexia and Obesity in Various Subreddit Forums

Navraj Narula and Ratheet Pandya
[nnn2112, rp2707]@columbia.edu
Computer Science Department, Columbia University
New York, NY 10027 USA

Abstract

Social media websites and online platforms offer abundant data regarding user behavior. In our paper, we focus on Reddit—a rather unconventional platform that aggregates news posts, collects web content, and hosts discussions. Subreddits, which are forums dedicated to specific topics on the site, of particular concern to us are those that include language related to anorexia nervosa and obesity. We examine comments under such posts and extract features that may be useful for determining whether or not a user suffers from one of the aforementioned eating disorders. A prediction experiment on the text of comments unrelated to these conditions, but made by users who have posted about eating disorders, was conducted as well. This was used to determine salient features in text written by users suffering from these disorders. We discuss our results in conjunction with prior work surrounding topics related to eating disorder analysis in text.

1 Introduction

According to the National Association of Anorexia and Associated Eating Disorders, at least 30 million people suffer from an eating disorder in the United States. Among this population, a death occurs every 62 minutes as a result of poor nutrition, weight imbalance, or

feelings tending towards depression or suicide [1]. Although the dialogue surrounding eating disorders is active, research pertaining to the subject is largely centered around biological processes. With populations now expressing themselves in both clinical settings and online forums, it becomes crucial to incorporate language examination into research surrounding eating disorders. Our study focuses on two contrasting eating disorders: anorexia nervosa and obesity.

Anorexia nervosa is a psychological illness characterized by low body weight, fear of weight gain, and a distorted perception of the self. It has the highest mortality rate of any mental illness [2]. Obesity is defined by an excess of adipose tissue. This excess leads to increasing circulation of fatty acids, which can result in insulin resistance that can then lead to type II diabetes. [3].

Unlike anorexia, poor self-image is not necessarily the cause of obesity; however, it could be associated with the condition as it develops. Over 35 percent of U.S. adults are at risk for obesity, while anorexia affects nine percent of the country's population [4,1].

In our paper, we examine the language of anorexia nervosa and obesity in parallel. Our experiments involve a preliminary analysis of

the Reddit dataset. Following that, we extract Linguistic Inquiry Word Count (LIWC) features from the body of each subreddit and perform correlation experiments in efforts to determine dominant personality traits [10] present in anorexia nervosa and obesity. We also conduct a prediction experiment in which we attempt to predict whether an author has posted about anorexia, obesity, or neither topic in the past based on the text in the subreddit comment itself. This measure is conducted in efforts to determine the potential for predictive diagnoses of anorexia nervosa and obesity based on online forums.

2 Prior Work

Much of the research regarding the language of eating disorders falls into two major areas: survey-based analysis and predictive medicine.

In Sackville et al.'s analysis of attention women with eating disorders have towards food, body weight, and shape-related stimuli, participants were required to undergo multiple measures before partaking in a Stroop Task that included body-sensitive words [5]. Some of these assessments included an Eating Disorders Screening Form, the Eating Attitudes Test, the Beck Depression Inventory, and several others. Among all tests, participants were asked to rank themselves according to certain prompts. For example, a prompt in the Eating Attitudes Test might be "I feel extremely guilty after eating" and participants can select either "always," "usually," "often," "sometimes," "rarely," or "never."

Studies in this area tend to involve a small number of participants and the participant pool

is somewhat homogenous. In Wolf et. al.'s study of cognitive style in journal entries written by patients facing eating disorders, only 11 people participated in the study. Out of these 11 people, only two people were non-Caucasian and only one of them was non-female. The main result of the study indicated that journals displayed the highest rates of self-related words, negative emotion words, and the lowest rate of positive words [6].

In particular, in-person studies seemed to be more common in literature related to anorexia. The largest study pool came from Davies et al.'s study related to the analysis of verbal expression of emotion in speech from people who have anorexia or bulimia. The specific participants themselves included 42 people with anorexia, 26 people with bulimia, and 20 people with neither eating disorder who were all video-taped talking about discrete emotional experiences, after being provided a prompt [7]. The small participant-size is likely to be expected as exploring natural language processing applications in eating disorders is a fairly new field.

In studies pertaining to obesity, data was often collected via online sources with larger sample sizes. Harris et al.'s study of communication of childhood obesity on Twitter included a sample size of 1100 tweets originating from 576 Twitter users [9]. Abbar et al.'s study regarding food consumption using demographic data on Twitter considered 210K users [8].

Regarding methods of analysis, use of LIWC is quite common [10]. Davies et. al. specifically used LIWC to analyze verbal expressions of emotion in participants who had anorexia,

bulimia, or no eating disorder at all. During this experimental task, participants were asked to respond to a computer asking them about neutral topics (i.e. “please tell me what the weather has been like” or “please tell me about something which has made you feel sad”). The responses to these questions were then analyzed using LIWC, which indicated that people who had anorexia used less positive words than those who did not. Davies et al. predicted that those who had bulimia would show “some impairment in expression of emotion;” however, the bulimia group was “indistinguishable” from the control group in terms of LIWC variables such as total word count and the relative use of positive words [7].

Abbar et al.’s paper on food consumption in Twitter also incorporates LIWC. A binary vector of LIWC categories were computed for each user of food names mentioned in their tweets [8]. The aforementioned study, by Wolf et. al., regarding journaling also made use of LIWC categories to describe writing styles of patients with eating disorders [6].

In the domain of predictive medicine, Ware et al. used a research corpus consisting of 700 anonymized patient discharge summaries that were labeled with the presence or absence of 16 different patient conditions, like diabetes or obesity, and attempted to predict these concepts for an NLP Challenge with a focus on obesity. They used synonym lists for medication names and used simple regular expression matching against each document for these features [4].

2.1 Hypothesis

Prior work indicates that eating disorders have a strong relation to negative self-image and self-worth. We hypothesize that those with eating disorders under the umbrella of anorexia and

obesity are likely to express themselves negatively on social media forums as well. Since anorexia induces more of a concern in regards to body image, we expect to see the use of negative words, such as words that may be associated with negative connotations in LIWC, to be more prevalent in anorexia-related posts. Fewer words, terse phrases, and self-references when speaking about weight may also be more common in anorexia. We hypothesize that obesity may exhibit similar features; however, results may be centered more around bodily functions and processes than self-evaluation of weight.

3 Dataset

The dataset used for our experiments consists of one month’s worth of publicly available comments from the online forum, Reddit [12]. This amounts to roughly 53.8 million JSON objects representing user comments before preprocessing. The full dataset was originally assembled in July 2015.

Each JSON object consists of metadata such as who posted the comment (`author`), the text of the comment (`body`), the name and ID of the topic (`subreddit`, `subreddit_id`), and additional metadata such as a comment’s `score` (based on user votes), when it was retrieved, etc.

Given the size of the corpus and computational resource constraints, we preprocessed the dataset to a more manageable size.

3.1 Preprocessing

Preprocessing involved filtering and transforming the full dataset into a smaller

dataset that is easier to use for our experiments. This amounted to extracting comments salient to our task and transforming the data from its JSON representation to CSV.

We used the `pandas` package for reading in the raw JSON objects and transforming them into CSVs for further preprocessing [13]. We also used the Natural Language Toolkit (`nltk`) for light preprocessing to tokenize the text of each comment for determining which comments to retain and which to discard [14]. This stage consisted of discarding comments with fewer than 10 tokens, since these comments were considered too small to contain information relevant to our experiments, and only including the `author`, `body`, `subreddit`, `subreddit_id`, and `score` columns.

Once the data was represented as a CSV file, we did further filtering to obtain comments relating to anorexia or obesity by obtaining synonym sets for ‘anorexia’ and ‘obesity’ from WordNet, and creating separate CSV files for those comments which contain any of the synonyms for each of these topics [15]. The resulting CSVs contained 1557 instances related to anorexia and 7465 instances related to obesity. This allowed us to create separate `pandas` data frames for each topic, which made it easier to join data frames and filter out data pertaining to either, both, or neither topic, as well as for performing analysis on each topic separately.

The code for this initial preprocessing step can be found in the “Data Preprocessing” notebook in our project’s GitHub repository [16].

3.2 Anonymizing the Data

In order to keep user names of comment authors anonymous for the purposes of our experiments, we transformed each comment author using an MD5 hash so that one could not easily determine which author posted about which topic. Note that this is an imperfect way of anonymizing the data since there may be other confounding factors, e.g. references to author names in the comment text, that make it difficult to anonymize comments entirely.

4 Experiments

The goal of our research is not only to highlight semantics used when discussing anorexia and obesity, but also to determine whether or not a diagnosis of anorexia and obesity is possible when analyzing online text present in social media. Our experiments involve a preliminary exploration of the dataset, LIWC feature extraction, and a prediction task.

4.1 Preliminary Analysis

In our preliminary analysis of online user data, we consider various attributes of interest in both datasets, covering posts related to anorexia and covering posts related to obesity. One particular attribute of interest is `subreddit`. To determine which subreddit most users posted to, we translated our datasets into `pandas` dataframes and created a dictionary containing frequency counts of which subreddits most users posted to. The dictionaries for both anorexia and obesity-related posts are visible in our “Analysis” notebook in our project’s GitHub repository [16].

Dialogue concerning both anorexia and obesity were largely centered on the *fatpeoplehate* subreddit, with 27.7 percent of posts related to anorexia and 25.8 percent of posts related to obesity. In light of Reddit’s new anti-harassment policy released in 2010, this particular subreddit—which previously had over 5000 subscribers—was banned due to its mission of posting pictures of overweight people for the “sole purpose of ridicule” [17].

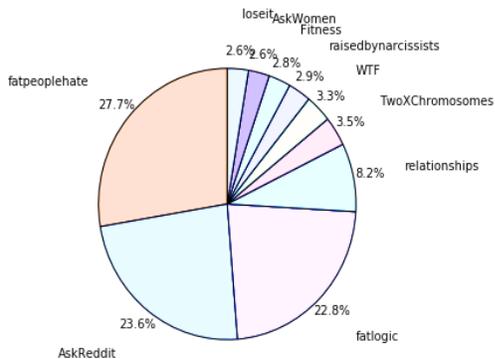
AskReddit, a general question-and-answering board on the platform for any topic, and *fatlogic*, a forum that encourages discussion of weight excluding scientific facts [18], follow closely after *fatpeoplehate* in serving as the most common forum posted to for both eating disorders. *AskReddit* hosted 23.6 percent of anorexia-related posts and 25.5 percent of obesity-related posts. *fatlogic* contained 22.8 percent of anorexia-related posts and 23.3 percent of obesity-related posts. These results are made more explicit in the diagrams below.

Commonality of subreddits posted to differ among anorexia and obesity-related in the following most common subreddits indicated on the pie charts.

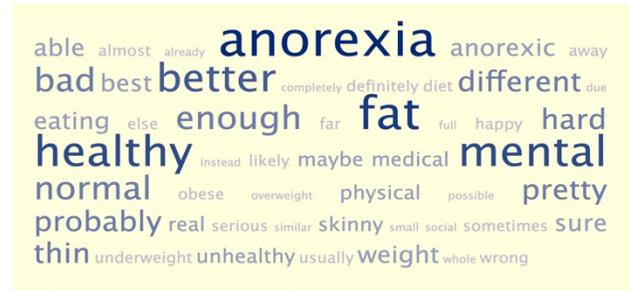
Following *fatpeoplehate*, *AskReddit*, and *fatlogic*, posts related to anorexia center on the *relationships* subreddit, a space to discuss any topic concerning romance, and the *TwoXChromosomes* subreddit, a forum intended to drive the “women perspective” among any topics [19]. Posts related to obesity were directed more present in the *funny* subreddit, which emphasizes humor.

Another attribute of interest was the *body* of each post. By means of natural language processing and heavy use of the nltk library, we extracted the most common words present in subreddits related to both anorexia and obesity. The word clouds for each are indicated in the figures below, generated by the TagCrowd system [20]:

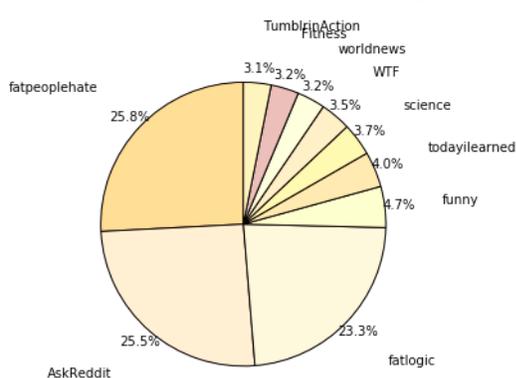
Common Subreddits for Anorexia Posts



Common Words in Anorexia Posts



Common Subreddits for Obesity Posts



Common Words in Obesity Posts



Prior to generating each word cloud, stopwords were removed. When irrelevant words still appeared in our results, we added part-of-speech tags to each word. This helped us to determine which part-of-speech categories contained words that delivered more telling information regarding anorexia and obesity. We found that adjectives and adverbs revealed more tangible information; thus, these are the part-of-speech tags most commonly associated with the words in our cloud. Manual preprocessing was done to remove further irrelevant words. Words were considered in isolation, and not in context.

4.2 LIWC analysis

LIWC is a commonly used text analysis program, which calculates the degree to which people may use different categories of words in text [21]. The `body` attribute for each of our datasets was fed into a LIWC program, which calculated the percentage of words in a given post that would fall into each category. In our analysis, we chose to omit consideration of dictionary words. After translating our LIWC results into a `pandas` dataframe, we determined the most common LIWC categories present among anorexia and obesity-related posts. This task was accomplished by calculating which category contained the highest percentage per instance in the dataframe.

After determining the most dominant LIWC categories for each eating disorder, we performed correlation experiments regarding the most dominant categories using R. These experiments are visible in our repository in the main folder [16], but are made more explicit in our results section.

4.3 Prediction Task

In the prediction task, the goal was to determine whether it is possible to predict, solely based on the text of a comment, whether the comment author had previously posted about anorexia, obesity, or neither. The motivation for this task was to find whether there are aspects of comment text from users who have posted about these topics that might be discriminative or specific to the way those users express themselves online.

The anorexia and obesity CSVs were used to create balanced training and test data sets with 5000 training records and 5000 test records for each class in the set {anorexia, obesity, neither} for a total of 30K records. In particular, this data consisted of comments made by users who have posted about anorexia, obesity, or neither topic, but the comments themselves did not pertain to these topics specifically. In other words, the anorexia and obesity CSVs were used to extract a list of users which fell into each category, and comments made by these users on arbitrary topics not relating to eating disorders were extracted for the prediction task. In this way, we hoped to establish whether comment text not relating to an eating disorder might still signal whether the user has posted about the eating disorder previously, suggesting a path for predictive diagnoses.

The training set was used to train an `scikit-learn` pipeline consisting of unigram and bigram counts, a term frequency-inverse document frequency (tf-idf) vectorizer for feature generation, and a linear model that uses stochastic gradient descent to optimize the loss function [17].

Use of tf-idf for the prediction task was done with each comment modeled as a document, and tokens in each comment corresponding to terms. We performed a grid search over the following hyperparameters:

- max document frequency: a threshold value in $[0, 1.0]$ that describes which terms to discard. For example, a max document frequency of 0.9 means to discard terms which appear in $> 90\%$ of documents, such as stop words like ‘the’
- ngram range: whether to use only unigrams or both unigrams and bigrams
- alpha: a coefficient that multiplies the regularization term in the loss function
- penalty: which form of regularization to use in the loss function (L2-norm or elasticnet)

The best hyperparameters found were a max document frequency of 0.75 (meaning to discard terms which appear in more than $\frac{3}{4}$ of the comments), using both unigrams and bigrams, an alpha parameter of 1×10^{-5} , and using L2-norm for regularization.

The code for prediction can be found in the “Prediction” notebook in our project’s GitHub repository [16]. The sklearn pipeline code was largely adapted from [22, 23].

5 Results

We discuss our results on analyzing correlations within LIWC categories and performing the prediction task.

5.1 LIWC analysis

The most dominant summary variable in LICW categories among both anorexia and obesity-

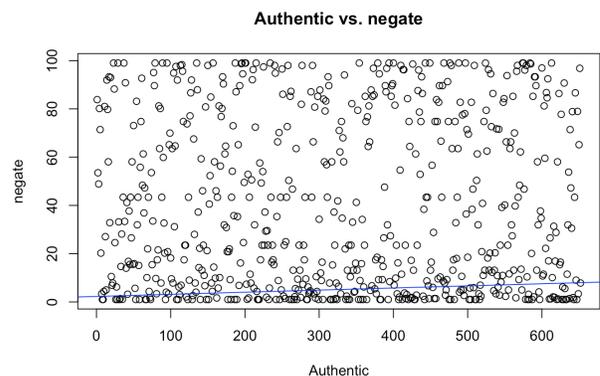
related posts was Clout, and following that, Analytic. Clout suggests that the author is highly confident in their speech and Analytic reflects logical thinking [21].

Posts related to anorexia also demonstrated more Authenticity, or more honesty, than posts related to obesity. The latter was associated more so with Tone, or an upbeat connotation in speech [21].

Based on the results regarding these summary variables and other dominant LIWC categories in the text, correlation experiments were conducted in R to determine whether or not a correlation exists among LIWC categories when considering p-values to be significant if less than or equal to 0.05.

We found that strong correlations exist among these variables for anorexia-related posts (Negate and Authentic are graphed below):

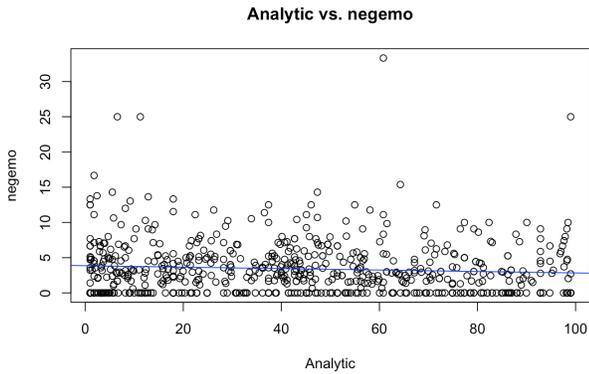
- Clout and pronoun
- Negate and Authentic
- Authentic and Analytic



In regards to obesity-related posts, we found that the strongest correlations included (Analytic and negemo are graphed below):

- Analytic and negemo

- Health and Analytic
- Clout and AllPunc



5.2 Prediction Task

The precision, recall, and F1 scores from the prediction task on the test data set are summarized in Table 1.

| | Precision | Recall | F1-Score |
|----------|-----------|--------|----------|
| anorexia | 0.63 | 0.92 | 0.75 |
| obesity | 0.52 | 0.38 | 0.44 |
| neither | 0.62 | 0.51 | 0.56 |
| Average | 0.59 | 0.60 | 0.58 |

Table 1: Prediction Task result summary

The recall and F1-scores of comments associated with users who have posted about anorexia stand out, while comments from users who have posted about obesity have lower classification precision and recall as compared to comments from users that have posted about neither topic, which we use as a baseline.

Our interpretation of these results is that users who have posted about anorexia in the past use language that is slightly more discriminative than users who have posted about neither area,

all else being equal. The reason we say this is that the strong recall means that of those comments posted by users who have previously posted about anorexia, 92% of them were correctly identified as having been posted by those users. In other words, the classifier was able to predict that the author of a given comment not related to an eating disorder was written by a user that has posted about anorexia in the past, 92% of the time. It is possible that a confounding factor, such as a correlation with anorexia-related comments not filtered out during preprocessing for the prediction task, is responsible for the unexpectedly high recall. Another source of noise is the potential that users who post about anorexia may suffer from this condition more than those who post about obesity, and that these users may use special “throwaway” accounts on Reddit to post about these or related issues, in order to avoid being stigmatized online when posting about other topics. More investigation in this area is needed.

The precision and recall for comments made by users who have posted about obesity is also interesting in that these metrics were worse than for comments made by users who have posted about neither topic (i.e., the baseline). We interpret this to mean that users who have posted about obesity may not have any particular specificity to their language so that the task of predicting this condition solely from comments on social media forums requires more work.

6 Discussion and Future Work

The Reddit data set is rich with information on social media behavior that could be used for interesting analysis and classification tasks, like

those we have done in this project. Accompanying this richness is a fair amount of noise, since literally anyone can post about anything on Reddit and it takes some work to filter the data set down to comments of interest for experiments.

LIWC results mirrored both our prior research and hypothesis, which stemmed from Davies et al.'s findings that suggested that those with eating disorders (i.e. particularly, anorexia) use fewer positive words. Our results indicate that negative words and words related to negative emotion were not only heavily present in both eating disorders, but also indicated correlation among other variables as well. In anorexia, negative words were highly correlated with the Authentic variable. In obesity, negative emotion was highly correlated with logical thinking. The latter was not to be expected since dialogue concerning eating disorders is more so tied to illogical reasoning; however, this may simply just indicate a tendency to justify a negative emotion. The same could be said for Clout, particularly in terms of anorexia.

In regards to the prediction task, predicting users in the anorexia class performed the best with a strong recall, suggesting that there are features of the language of users who have posted about anorexia that is more specific than users who post about other topics. On the other hand, predicting users in the obesity class performed the worst, suggesting little to no discriminative features in social media text for these users.

It is possible that these results point to a bias in the data set rather than an objective property of social media text written by users who suffer

from anorexia or obesity. This requires further investigation. For example, in order to rule out the potential that users might be using throwaway accounts to discuss eating disorders or related topics, one would need to do some analysis on the usernames of comment authors in relation to which topics they post about, which we avoided for the sake of protecting privacy.

To summarize, while Reddit data can be a great source of information about how individuals express themselves regarding eating disorders, it is important to carefully filter and preprocess the data set to manage noise inherent in user comments. Future work in this area should address these issues, and perhaps build classifiers using LIWC features to compare whether feature generation using LIWC improves precision/recall metrics over those of raw text.

References

- [1] "[Eating Disorder Statistics](#)," National Association of Anorexia and Associated Eating Disorders (ANAD). Accessed April 2017.
- [2] Smink, F. E., van Hoeken, D., & Hoek, H. W. (2012). Epidemiology of eating disorders: Incidence, prevalence and mortality rates. *Current Psychiatry Reports*, 14(4), 406-414.
- [3] "[What is Obesity?](#)" The Obesity Society. Accessed April 2017.
- [4] "[How many people are affected by/at risk for obesity & overweight?](#)" Eunice Kennedy Shriver National Institute of Child Health and Human Development. Accessed April 2017.

- [5] Sackville, T., Schotte, D. E., Touyze, S.W., Griffiths, R., and Beumont, P.J.V. “[Conscious and preconscious processing of food, body weight and shape, and emotion-related words in women with anorexia nervosa.](#)” International Journal of Eating Disorders. January 1998.
- [6] Wolf, M., Sedway, J., Bulik, C. M., and Kordy, H. “[Linguistic analyses of natural written language: Unobtrusive assessment of cognitive style in eating disorders.](#)” International Journal of Eating Disorders.
- [7] Davies, H., Swan, N., Schmidt, U. and Tchanturia, K. (2012), “[An experimental investigation of verbal expression of emotion in anorexia and bulimia nervosa.](#)” Eur. Eat. Disorders Rev., 20: 476–483. doi:10.1002/erv.1157
- [8] Abbar, S., Mejova, Y., Weber, I. “[You Tweet What You Eat: Studying Food Consumption Through Twitter.](#)” Qatar Computing Research Institute, 2015.
- [9] Harris, J.K., Moreland-Russell, S., Tabak, R.G., Ruhr, L.R., Maier, R.C. “[Communication about childhood obesity on Twitter.](#)” American Journal of Public Health. July 2014.
- [10] Y. R. Tausczik and J. W. Pennebaker. “[The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods.](#)” Journal of language and social psychology, 29(1):24--54, 2010.
- [11] Ware H., Mullett C., Jagannathan V. “[Natural Language Processing Framework to Assess Clinical Conditions.](#)” Journal of the American Medical Informatics Association. 2009;16(4):585–89.
- [12] The Reddit /r/datasets online forum, 2015, https://www.reddit.com/r/datasets/comments/3bxlg7/i_have_every_publicly_available_reddit_comment/. Accessed March 2017.
- [13] Wes McKinney. Data Structures for Statistical Computing in Python, Proceedings of the 9th Python in Science Conference, 51-56 (2010) ([publisher link](#))
- [14] Edward Loper and Steven Bird. 2002. “[NLTK: The Natural Language Toolkit.](#)” In Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics, pp. 62–69. Somerset, NJ: Association for Computational Linguistics. <http://arXiv.org/abs/cs/0205028>.
- [15] Princeton University "About WordNet." WordNet. Princeton University. 2010. <http://wordnet.princeton.edu>.
- [16] Narula, N. and Pandya, R. Subreddit Analysis on Eating Disorders, 2017, GitHub repository, <https://github.com/navierula/Subreddit-Analysis-on-Eating-Disorders/>. Accessed April 2017.
- [17] Robertson, A. “Reddit bans 'Fat People Hate' and other subreddits under new harassment rules,” The Verge, June 2015. URL: <https://www.theverge.com/2015/6/10/8761763/reddit-harassment-ban-fat-people-hate-subreddit>
- [18] “[/r/fatlogic](#),” Accessed May 2017.

- [19] “[/r/TwoXChromosomes](#),” Access May 2017.
- [20] “TagCrowd,” URL: <http://tagcrowd.com>. Accessed May 2017.
- [21] Pennebaker, J.W., Booth, R.J., Boyd, R.L., Francis. M.E. “[Linguistic Inquiry and Word Count: LIWC2015](#),” Accessed May 2017.
- [22] [Scikit-learn: Machine Learning in Python](#), Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.
- [23] Scikit Learn “Sample pipeline for text feature extraction and evaluation.” http://scikit-learn.org/stable/auto_examples/model_selection/grid_search_text_feature_extraction.html. Accessed April 2017.
- [24] Scikit Learn “Parameter estimation using grid search with cross-validation.” http://scikit-learn.org/stable/auto_examples/model_selection/grid_search_digits.html. Accessed April 2017.