

TEXT MINING MENTAL HEALTH FORUMS – LEARNING FROM USER EXPERIENCES

Research in Progress

Davcheva, Elena, Technical University Darmstadt, Darmstadt, Germany, davcheva@ise.tu-darmstadt.de

Abstract

Mental healthcare today represents a serious global challenge with not enough resources to allow for adequate patient support. As a result, many turn to the Internet for help, where mental health forums have become a rich resource of experiences shared by millions of users. This study applies aspect-based sentiment analysis on mental health forum posts in order to examine user sentiment regarding different mental health treatments. We shed light into the practices used by affected individuals to cope with mental issues and generate possible treatment recommendations.

Keywords: text mining, mental health, sentiment analysis, big data

1 Introduction

Mental health today has become a global issue. The World Health Organization reports that around a quarter of all years lived in disability is owed to mental disorders such as depression, bipolar disorder, or substance abuse, to name a few. Each year, around 800,000 persons in the world commit suicide, making it the second leading cause of death among 15-29 year-old individuals globally (Suicide data, 2017). Regardless of these numbers, with 60 countries reporting the availability of less than one psychiatrist per 100,000 population (Mathers, 2008), the number of mental healthcare professionals is not nearly enough in order to adequately address the needs of all those seeking help (Kauer et. al., 2014). Of those affected, many are unable to afford professional mental healthcare, as it also happens to be the costliest medical condition to treat (Saleem et. al, 2012). Professional psychological help usually requires patients to undergo expensive and time-consuming clinical tests and in-person interviews (Weathers et. al, 2001). In the US alone, providing mental healthcare costs \$200 billion a year (Roehrig, 2016), thus causing also a burden on healthcare planners and providers. Finally, regardless of their personal financial situation, many affected individuals would not even consider professional mental healthcare because of fear of stigma and peer rejection (Crabtree et. al., 2010).

As a result of the state of the global mental healthcare system, more and more people turn to the Internet in a bid to get the help they need rapidly, affordably, and anonymously. Online forums devoted to mental health discussions are growing. In the US alone, 80% of those with Internet access use it to get health-related information, and 34% of those look up specifically personal stories from other users (Malmasi et. al., 2016). Forums are free of charge and open for anyone to join discussions, share experiences, provide answers, or ask their own questions. Users flock to forums thanks to the anonymity of the Internet, the ease of use, as well as the convenience of time and location independence (Kummervold et. al, 2002). Users feel that forum participation is a rewarding experience because it makes them feel as part of a group; they get the chance to share the burden of living with a certain condition, connect to others with similar experiences, and learn from others' successes and mistakes (Eysenbach et. Al, 2004). It is a place that demands low investment, but participation can prove very rewarding.

Health forums prove to be an obligatory source of information for mental health related questions, as even when people do visit a medical professional, they still participate in forum discussions nonethe-

less (Kummervold et. al, 2002). Even more importantly, the age distribution of forum participants suggests that as digital natives grow up, online forums will become even more relevant in the future (Kummervold et. al, 2002). Online tools such as the forums have the potential to become a tool to address the shortcomings of the mental healthcare system and the ever-growing number of people looking for help.

Past studies show that users are more honest and more prone to sharing personal stories online (Barak & Gluck-Ofri, 2007). The data within the forum discussions is therefore a valuable source of information to researchers wishing to understand more about people suffering from certain conditions.

2 Motivation and Design

This study makes use of user posts on online forums about mental health and tries to make sense of what users talk about – namely, how individuals suffering from different conditions express themselves about various topics that can have a positive influence on their condition, e.g. therapy, doctors, meditation, or sports. The goal of the study is to produce a cross-condition comparison of the sentiments expressed for these concepts. For that goal we employ aspect-based sentiment analysis based on linguistic modelling techniques for natural language.

Natural language data are both very valuable and difficult to process because of their inherent lack of structure and formality. Text mining and processing large volumes of forum posts requires constructing a sophisticated data processing pipeline which can identify complex grammar structures and word interactions within a specific language, as well as having the capacity to discern different emotional nuances in words and phrases (Saleem et. al, 2012). Such analyses are time-consuming and require much computing time and power. This paper combines the application of state-of-the-art NLP techniques onto a large dataset with a novel research question in order to advance our understanding of mental health experiences and provide recommendations to enhance future treatment approaches.

The rest of the paper is ordered as follows: Section 2 presents a detailed review of research into online mental health interactions, focusing specifically on the use of sentiment analysis and NLP techniques. Section 3 presents the methodology of the study and the data used. Section 4 presents the analysis results. Section 5 summarizes findings and limitations, as well as suggests next steps.

3 Literature Review

Research into the use of online mental health spaces represents a recent effort which is gaining more and more traction as the Internet becomes an important space for mental health information and communication. Even though participation in forums is not a replacement for therapy, the goal of therapy is to induce a positive change in behaviour, and there is already evidence that usage of online mental health aids such as forums, social media, or chatbots, leads to changes in behaviour. A few studies show that both online aids with and without medical professionals' participation can lead to comparatively effective results, particularly in the cases of alcoholism (Riper et. al., 2014), smoking (Aveyard et. al, 2012), anxiety (Cuijpers et. al., 2009), and PTSD (Kuester et. al., 2016).

In terms of research methods, researchers have frequently employed manual techniques as tools to analyze forum conversations, such as forum user surveys or discourse analysis on a small sample of user posts. Recently, automated text mining techniques such as sentiment analysis and NLP have also been applied. Thematically, research subjects vary from classifying post helpfulness to measuring and comparing sentiment, to identifying the presence of specific content in single posts. Previous studies find online communities generally helpful for various mental conditions. Johnsen et. al. (2002) use human readers to classify mental-health forum interactions as either helpful or unhelpful. While using human readers is regarded as a reliable analytical tool, human processing speed limits the amount of posts that can be processed in the analysis – in this case to only 102, which is why automated analysis is recently introduced as a way to take advantage of available big data sets. Specifically regarding sentiment analysis, even though it has been established as a reliable technique, there are still very few papers using the method in the context of online mental health forums. A fundamental effort by Nguyen et. al.

(2014) compares the sentiment expressed on depression forums with sentiment on non-depression forums, showing that individuals who are not depressed generally express themselves more positively. Thus the study directly supports the use of sentiment analysis (SA) as a viable tool in analysing mental health forums. SA has also been applied in pharmacovigilance, or the identification of adverse drug effects, by automatically identifying medications with negative opinions on social media (Korkontzelos et. al. 2016) – thus granting researchers the possibility to explore drug effects over a large and diverse population. The study is also an important step in using text mining to detect possible negative effects of treatments that might generally be considered safe. Further uses of SA in mental health forums include monitoring the influence of social media messages on potential behavioural changes. Namely, through the application of sentiment analysis, Cobb et. al. (2013) showed that positively discussing certain actions which lead to quit smoking will lead to social media users actually implementing these actions in real life, a testament to the influence that online user-to-user communication has on people’s life choices. Wang et. al. (2013) used SA as a classification tool to identify if a user has a certain condition – such studies may lead to future development of automated online diagnostic tools (Saleem et. Al, 2012).

Coppersmith et. al. (2015) show that the presence of positive emotion is not an indicator for the presence of a mental health condition. On the other hand, expressing a variety of negative emotions can be an indicator for all of the conditions examined in this study. Furthermore, through an analysis of social media messages, De Choudhury et. al. (2013) also show that the positive or negative affect of written language is one of the best language features to predict depression.

More recent studies have begun focusing on determining beneficial practices for specific conditions by applying content and topic analysis. Spijkerman et. al. (2016) demonstrated the benefits of meditation and mindfulness practices delivered online for people suffering from depression and anxiety by utilizing manual assessment. A similar study investigates the perception that cannabis can successfully treat ADHD – by using human coders to analyse natural text, it shows that 25% of ADHD users who have self-medicated with cannabis reported positive experiences (Mitchell et. al., 2016). As it can be seen, many of these efforts rely on manual assessment, which presupposes working with limited-size datasets. Automated NLP has also been applied in mental health forum analysis to research diverse topics such as the presence and effects of stigmatizing individuals with mental health conditions, efforts to automatically identify suicide ideations (with the goal of timely intervention and eventual prevention), topic modelling, and identifying specific emotions within user posts (Calvo et. al., 2017). Many of these efforts have focused on Twitter, where posts are limited to 140 characters. In terms of data size, the research field has yet to make proper use of the large amount of data available online. Johnsen et. al. (2002) used only a small sample of a one-month dataset to be processed and analysed by human readers. Mitchell et. al. (2016) uses only 55 threads and 401 posts altogether.

To the best of our knowledge, there exist no previous studies which have attempted to create a cross-condition comparison of potentially helpful treatments for mental health via aspect-based sentiment analysis. Previous studies also tend to focus on single conditions or specific features, while the goal of our study is to compare several features across several conditions.

4 Dataset and Methodology

We use aspect-based sentiment analysis (ABSA) to determine the sentiment forum users express on various aspects or concepts such as family or doctors, which can potentially help them on their path to healing. SA can be subject-dependent and subject-independent (Wang et. al., 2013). A subject-independent SA measures the sentiment of chunks of text, e.g. a sentence or a post. On the other hand, subject-dependent SA measures the sentiment expressed regarding a particular subject. The latter is also known as ABSA and is based on parsing natural text through linguistic NLP dependency parsers that match subject words with other words that directly describe or relay a quality of the subject words. For example, after parsing a simple sentence such as “These apples are green”, an NLP dependency parser would return the pair (apples, green), where “apples” is recognized as the subject, and “green” as the descriptive word.

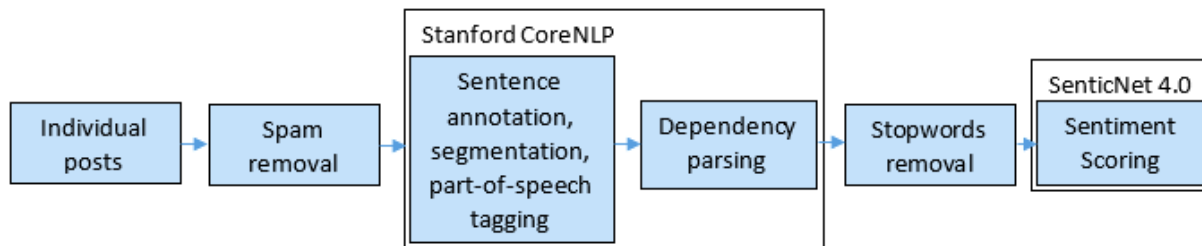


Figure 1. Analysis Procedure

In this study, our subject words are all words connected to a concept whose sentiment we want to measure, as seen in (Wang et. al., 2013) and other studies. The concepts we measure SA for are family, medications, therapy, pets, sports, and meditation. Each of the concepts is a sum of sentiments expressed for all words related to itself, e.g. the sentiment for family is the sum of sentiments for all family-related words, e.g. family, parents, siblings, mother, sister, son, etc. For the specific calculation of the sentiment score, we use the SenticNet 4.0 dictionary by following the approach of Taboada et. al. (2017). That is, apart from using word scores, we also account for intensification (e.g. “bad” vs. “very bad”) and negation (e.g. “good” vs. “not good”). Figure 1 describes the analysis process used in this paper.

The first step in the data processing is removing spam, or in our case – posts that have been quoted in later replies. Then the posts are tokenized (divided into words), segmented into sentences, and tagged with part-of-speech information, e.g. noun, verb, adjective. We use the Stanford CoreNLP models for the tasks of tokenization, part-of-speech tagging, sentence segmentation, and dependency parsing. Namely, CoreNLP provides an English-language syntactic dependency parser based on a recurrent neural networks model. The dependency parser is a crucial part of this analysis, as it models the grammatical structure of a sentence and provides information as to relationships between different words (Neural Networks Dependency Parser).

| Relation | CoreNLP Definition | Example Sentence |
|-------------------------|---|---|
| Adjectival modifier | any adjectival phrase that serves to modify the meaning of a noun | Sam eats red meat (meat, red) |
| Nominal modifier | nominal dependents of another noun or noun phrase that functionally correspond to an attribute or genitive complement | The Chair’s office (chair, office) |
| Nominal subject | a nominal which is the syntactic subject and the proto-agent of a clause | The baby is cute (cute, baby) |
| Open clausal complement | predicative or clausal complement | He says that you like to swim (like , swim) |

Table 1. Stanford CoreNLP Dependency Relations used in this study

The neural network models and rules are described in-depth in the paper by Chen and Manning (2014). In this way, for each concept and its related terms we can put together a sub-selection of relevant descriptive words. The sentiment of each of the 8 concepts examined in this paper is calculated as the average sentiment (from SenticNet) of all descriptive words for that concept as identified by the dependency parser; i.e., for each concept SC_i the sentiment is found by the formula $SC_i = \frac{\sum SC_j}{n}$, where n is the number of related terms to that concept, and SC_j is the sentiment score of an SC_i related term as provided by SenticNet. The syntactic relations we chose to focus on are presented in the table below with their respective CoreNLP documentation definition. These relations output word pairs where a subject word is described or modified by another word – these relations contain the most relevant word pairings for the needs of sentiment scoring.

We further focus our choice of words by taking into consideration only nouns, adjectives, and verbs within the above relations. This measure reduces noise by removing words in the above relations

which do not bear any descriptive qualities. As a result, sentiment accuracy will be improved. The SenticNet dictionary version 4.0 is what enables the aspect, or word-based sentiment scoring. In a nutshell, SenticNet is a list of 50,000 English-language words and their appropriate sentiment scores. We use SenticNet following the example of Wang et. al. (2013). Scores are adjusted if words are preceded by a negation. Three of the leading English-language mental health forums were scraped to create a combined set of 132,072 threads containing 1,155,403 individual posts across 12 conditions (depression, bipolar disorder (BD), anxiety and panic attacks, schizophrenia, attention deficit hyperactivity disorder (ADHD), Asperger's Syndrome, Borderline personality disorder (BPD), obsessive-compulsive disorder (OCD), post-traumatic stress disorder (PTSD), self-harming, substance abuse). The data were extracted in August 2017 and encompass all publicly available posts on the respective websites.

The forums have administrators and moderators whose task is to make sure conversations do not go off-topic; thus, we can be sure that in our research we are considering discussions relevant to each condition. Additionally, moderators remove offensive or damaging material (e.g. posts that encourage self-abusive behaviour). However, the role of moderators is not to provide advice, as the goal of a mental health forum is to be a place of discussion among the users, and not between a user and a medical professional, i.e. a forum is not meant or seen as a tool to replace established medical practices such as therapy.

5 Results

In terms of conditions, bipolar disorder forum users have expressed the highest sentiment across conditions (average 0.16), whereas autism forum users have the lowest average score of -0.04. The autism forum posts are the only ones to score an average negative sentiment, while simultaneously expressing the best sentiment for family and pets, while scoring meditation and spirituality the lowest. Clearly distinguishing family and pets as a positive presence suggests that emotional support by loved ones is very important for people with autism. This finding is backed by psychological research (Solomon, 2010) as well as recommendations of leading organizations as to the benefits of pets for autistic individuals (Autism and Pets, 2014).

| Concept \ Forum | Family | Sports and Exercise | Meditation and Spirituality | House Pets | Therapy | Medications | Medical Professionals |
|-----------------|--------|---------------------|-----------------------------|------------|---------|-------------|-----------------------|
| Anxiety | 0,086 | 0,077 | 0,184 | 0,098 | 0,225 | 0,075 | 0,123 |
| ADHD | 0,109 | 0,215 | -0,016 | 0,117 | 0,255 | 0,047 | 0,144 |
| Depression | 0,074 | 0,090 | 0,179 | -0,067 | 0,186 | 0,116 | 0,109 |
| Asperger's | 0,082 | 0,102 | 0,150 | -0,021 | 0,162 | 0,055 | 0,176 |
| OCD | 0,037 | 0,097 | 0,051 | 0,048 | 0,233 | 0,150 | 0,145 |
| BD | 0,088 | 0,154 | 0,250 | 0,109 | 0,263 | 0,071 | 0,163 |
| BPD | 0,098 | 0,137 | 0,242 | 0,060 | 0,180 | 0,071 | 0,148 |
| Self-harm | 0,078 | 0,186 | 0,006 | 0,092 | 0,263 | 0,117 | 0,149 |
| PTSD | 0,062 | 0,030 | 0,063 | -0,086 | 0,199 | 0,108 | 0,050 |
| Schizophrenia | 0,047 | 0,120 | 0,146 | 0,069 | 0,188 | 0,128 | 0,145 |
| Subs. Abuse | 0,064 | 0,320 | -0,063 | 0,090 | 0,287 | 0,083 | -0,023 |
| Autism | 0,176 | 0,018 | -0,355 | 0,149 | 0,0007 | -0,161 | -0,089 |

Table 2. Average sentiment per concept across mental health forums

Users in the substance abuse as well as self-harm forums have distinguished sports and therapy as the two most positive concepts. Substance abuse posters negatively score meditation, as well as medical

professionals other than therapists, which again draws similarities with self-harm, where meditation is the worst-performing concept.

Although all concepts score positively in the schizophrenia forums, therapy is most positively regarded, while family – the least. This may indicate that dealing with loved ones with this condition is not undesirable, but it remains challenging. Literature on the matter considers family relationships a necessary part of schizophrenia treatment, but a difficult and complex one (Motlova, 2007). Forum users express an extreme range of emotions regarding family – from a need for understanding and support to disinterest and rejection, however detailed research into the conversations is needed to extract signals as to ways in which a schizophrenic can better communicate with their closest ones while undergoing treatments. Regarding PTSD, therapy stands out as positive, while pets are the single negative concept. This finding contradicts a long-established practice to treat PTSD with pets such as dogs or horses (Altschuler, 1999). Looking more closely, the low sentiment score in our dataset is owed by the fact that many of the PTSD forum users report the loss or death of pets as one of several reasons that have triggered PTSD.

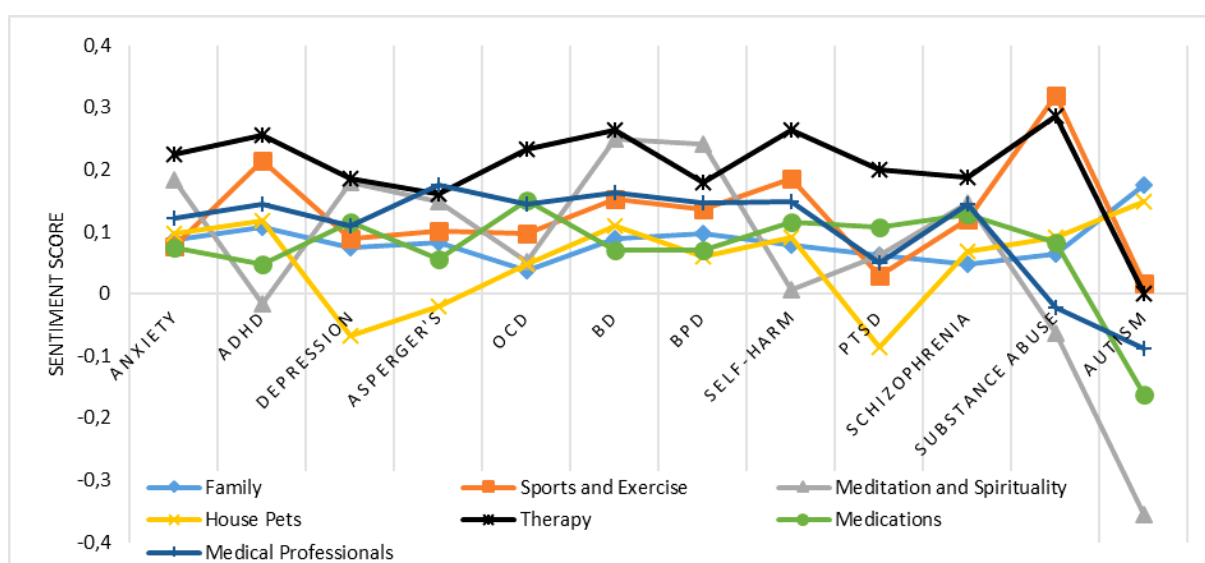


Figure 2. Sentiment Scores

BPD forum posters talk prominently most positively regarding meditation and spirituality. Meditative and spiritual healing practices vary widely between conditions, with BD and BPD individuals finding it the most helpful, the lowest scores coming from ADHD and substance abuse forums users. Despite some research efforts on the subject (Perich et. al., 2013), meditation is not yet established as a positive treatment for BD, and has not been considered for BPD. BD forum users highly recommend meditation to those seeking help and advice, characterizing it as “particularly helpful”, and report that meditation helps them in controlling the mood swings associated with the condition. BPD forums users report meditation to be a tool for avoiding psychotic states and dissociation. Most users report picking up meditation online, by using YouTube videos or mobile apps.

In terms of concepts, pets have the lowest sentiment score across all conditions, with an average score of 0,05 and a negative score in 3 out of the 12 conditions. PTSD and depression forum users often mention the loss of pets as a condition trigger. On the other hand, many in the Asperger’s forum share a genuine dislike of pets. Therapy is by a wide margin the most positively talked-about concept of all with an average sentiment of 0,2, and a highest sentiment score in 8 out of 12 conditions. This is a testament that professional one-on-one conversations are still the best treatment for most conditions, however this is not a rule of thumb, as 1/3 of forum users in all conditions surveyed feel more positive about another treatment. Therapy does score considerably lower within autism, having a neutral score of almost 0.0, despite many different types of therapy for autistic individuals, including speech, sensory, occupational, and cognitive-based therapy, to name a few (Blumberg et. al., 2016). This is backed

by the fact that about 13% of autistic individuals ever lose their diagnosis later in life, after going through the rigorous therapy plan (Blumberg et. al., 2016). Therefore, implementing therapy alternatives such as family support, house pets, and even physical activity, may be more beneficial for people with autism, as suggested also by our results. Some on the autism forum report being “hurt” by certain types of therapy, and many say that therapy was helpful but “not too much”.

Looking into sports and exercise, people with substance abuse, ADHD, or self-harm, stand to gain the most out of being physically active. For those with substance abuse this is the number one positive concept, and for ADHD individuals this is especially sensible, as being physically active plays into their hyperactivity (Lufi & Parish-Plass, 2011). There is not much in the literature as yet connecting substance abuse and self-harm with sports as a possible remedy for these behaviours.

Sports and meditation are concepts that score significantly high scores in certain conditions, which indicates that it is worthwhile to further investigate the possibility of incorporating these practices in the appropriate condition therapies also from a formal and professional point of view. On the other hand, getting pets appears undesirable across the 12 conditions, especially for Asperger’s forum users. Autism users are the exception when it comes to pets, where besides family, pets are the leading positive concept. It is also important to note that medications, albeit not negatively scored, are also not regarded with significantly high sentiment, consistently scoring significantly lower than therapy. Medications have the lowest sentiment among concepts within anxiety and BD forums, suggesting that these conditions gain the least advantage from using medication. Anxiety forum users report having mild anxiety even when using medications. An exception in this case is only the OCD forum, where medications are the second most positive concept.

Furthermore, in order to reveal the similarities between scores of different conditions, we conducted a two-tailed Pearson correlation analysis. Anxiety is strongly correlated with BD (Pearson 0.923, $p < 0.01$) and BPD (Pearson 0.771, $p < 0.05$). These findings are supported by psychological literature, as anxiety has high comorbidity with both conditions (Keller, 2006) (Zanarini et. al., 1998). Furthermore, ADHD is correlated to self-harming behavior (Pearson 0.922, $p < 0.01$) as well as substance abuse (Pearson 0.826, $p < 0.05$). The link between these two conditions and ADHD has also been well-documented (Wilens, 2004). This suggests that these conditions may gain similarly positive or negative results from the same treatments.

It is interesting to note that self-harm and BPD – two conditions usually associated with each other (Chapman et al, 2005), are strongly uncorrelated in our analysis. The difference is owed to the sentiment expressed on meditation and spirituality. In this case this may indicate that even though self-harm and BPD have a high comorbidity, different concepts or treatments may still elicit different responses.

5.1 Method Evaluation

Evaluating the accuracy of dictionary-based sentiment analysis requires evaluating the sub-tasks of aspect term extraction as well as aspect term sentiment evaluation. Both task results were evaluated using the standard metrics precision and recall (Salas-Zárte et. al., 2017). The system-generated aspect pairs and scores are checked against a subset of 2500 human-annotated pairs from the forum data. In the first sub-task, the goal is to make sure that aspect terms are related to and relevant for the subject term. In the second task, even though the SenticNet dictionary comes with word sentiment scores, it is necessary to check whether the aspect term scores make sense in the context of the subject term evaluated as well as the forums; for example, the pair (therapy, continued) was scored with a -0,04, while in context it had a more positive meaning.

The precision measure for aspect term extraction was 71.83%, and recall was 70.48%, whereas for sentiment scoring it was 78.12% for precision and 75.49% for recall. The lower measures regarding aspect term extraction as a sub-task signal that rules for extracting descriptive aspect terms specifically meant for a SA use must be narrowed down more precisely than those presented in our study. Nevertheless, the measures are encouraging when compared to other ABSA accuracy scores and render reliability to the study results (Da Silva et. al., 2014).

| | Anxiety | ADHD | Depression | Asperger's | OCD | BD | BPD | Self-harm | PTSD | Schizophrenia | Subs. Abuse | Autism |
|---------------|---------|--------|------------|------------|-------|--------|--------|-----------|-------|---------------|-------------|--------|
| Anxiety | 1 | ,157 | ,625 | ,618 | ,473 | ,923** | ,771* | ,240 | ,568 | ,716 | ,027 | -,380 |
| ADHD | ,157 | 1 | -,002 | ,194 | ,559 | ,211 | -,072 | ,922** | ,256 | ,247 | ,826* | ,568 |
| Depression | ,625 | -,002 | 1 | ,850* | ,533 | ,653 | ,766* | ,210 | ,866* | ,769* | ,032 | -,661 |
| Asperger's | ,618 | ,194 | ,850* | 1 | ,479 | ,714 | ,808* | ,284 | ,655 | ,723 | -,048 | -,495 |
| OCD | ,473 | ,559 | ,533 | ,479 | 1 | ,382 | ,160 | ,823* | ,756* | ,799* | ,449 | -,145 |
| BD | ,923** | ,211 | ,653 | ,714 | ,382 | 1 | ,907** | ,243 | ,458 | ,758* | ,116 | -,463 |
| BPD | ,771* | -,072 | ,766* | ,808* | ,160 | ,907** | 1 | -,043 | ,430 | ,657 | -,124 | -,658 |
| Self-harm | ,240 | ,922** | ,210 | ,284 | ,823* | ,243 | -,043 | 1 | ,499 | ,498 | ,802* | ,330 |
| PTSD | ,568 | ,256 | ,866* | ,655 | ,756* | ,458 | ,430 | ,499 | 1 | ,687 | ,278 | -,312 |
| Schizophrenia | ,716 | ,247 | ,769* | ,723 | ,799* | ,758* | ,657 | ,498 | ,687 | 1 | ,197 | -,622 |
| Subs. Abuse | ,027 | ,826* | ,032 | -,048 | ,449 | ,116 | -,124 | ,802* | ,278 | ,197 | 1 | ,436 |
| Autism | -,380 | ,568 | -,661 | -,495 | -,145 | -,463 | -,658 | ,330 | -,312 | -,622 | ,436 | 1 |

Table 3. Correlation Analysis (** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed))

$$Precision = \frac{True\ Positives}{True\ Positives - False\ Positives} \quad Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Figure 3. Precision and Recall Formulas

6 Conclusion

Mental health forums are a vast data source of information that formal psychology has not yet tapped into. This paper shows how natural text processing of large datasets can automate and accelerate data collection, processing, and insight generation from millions of posts by tens of thousands of individuals. Being able to back up many of the findings of this study with previous findings in psychological literature and in a few cases also practice, demonstrates how and why text mining of large public datasets is a viable and dependable tool that can bring about a veritable change in how we research mental health, and how we approach the question of improving and administering therapy and treatments.

The study singles out meditation and spirituality practices as well as sports and exercise as helpful practices in a variety of conditions such as BD and BPD. Although they are found to be positive by forum users, these practices have yet to become an established part of formal treatments. The results point to shortcomings of therapy in autism. As next steps the authors will take a deeper look into how exactly these concepts are being discussed within the forum data, in order to provide precise and detailed answers as to why sentiment scores appear to be what they are in this study. We will furthermore examine user-related variables such as location, mental health expertise and similar, in order to better understand the progression of sentiment. Finally, since autism is the most complex condition we look into in this study, there is much space to conduct a separate study focused only on this condition.

Although users express themselves positively for some concepts tested in this study, the current study does not measure exactly how effective users found these concepts to be in regards to their specific condition. More detailed linguistic analysis needs to follow in order to assess degree of effectiveness, as well as exact symptoms alleviated by each concept. Finally, this study raises the need to investigate the transfer of forum advice into the real world, specifically in terms of trust issues that may arise in an impersonal online environment (Benlian & Hess, 2011), as well as in terms of appropriate interface tools to aid users in finding the information that is pertinent to them (Benlian, 2015).

References

- Altschuler, E. L. (1999). Pet-facilitated therapy for posttraumatic stress disorder. *Annals of Clinical Psychiatry*, 11(1), 29-30.
- Autism and Pets: More Evidence of Social Benefits. April 15, 2014. Retrieved from <https://www.autismspeaks.org/science/science-news/autism-and-pets-more-evidence-social-benefits>
- Aveyard, P., et al. "Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: a systematic review and network meta-analysis." *Health technology assessment*, 16.38 (2012).
- Barak, A., & Gluck-Ofri, O. (2007). Degree and reciprocity of self-disclosure in online forums. *Cyber Psychology & Behavior*, 10(3), 407-417.
- Bell, V. (2007). Online information, extreme communities and internet therapy: Is the internet good for our mental health?, *Journal of mental health*, 16(4), 445-457.
- Benlian, A. (2015). Web personalization cues and their differential effects on user assessments of website value. *Journal of Management Information Systems*, 32(1), 225-260.
- Benlian, A., & Hess, T. (2011). The signaling role of IT features in influencing trust and participation in online communities, *International Journal of Electronic Commerce*, 15(4), 7-56.
- Blumberg, S. J., Zablotzky, B., Avila, R. M., Colpe, L. J., Pringle, B. A., & Kogan, M. D. (2016). Diagnosis Lost: Differences between Children who Had and who Currently Have an Autism Spectrum Disorder Diagnosis. *Autism : The International Journal of Research and Practice*, 20(7), 783-795. <http://doi.org/10.1177/1362361315607724>
- Calvo, R. A., Milne, D. N., Hussain, M. S., & Christensen, H. (2017). Natural language processing in mental health applications using non-clinical texts. *Natural Language Engineering*, 1-37.
- Chapman, A. L., Specht, M. W., & Cellucci, T. (2005). Borderline personality disorder and deliberate self-harm: does experiential avoidance play a role?, *Suicide and Life-Threatening Behavior* 35, no. 4 (2005): 388-399.
- Cobb, N. K., Mays, D., & Graham, A. L. (2013). Sentiment analysis to determine the impact of online messages on smokers' choices to use varenicline. *Journal of the National Cancer Institute Monographs*, 2013(47), 224-230.
- Crabtree, J. W., Haslam, S. A., Postmes, T., & Haslam, C. (2010). Mental health support groups, stigma, and self-esteem: positive and negative implications of group identification. *Journal of Social Issues*, 66(3), 553-569.
- Cuijpers P, Marks I.M., van Straten A. et al. (2009). Computer-Aided Psychotherapy for Anxiety Disorders: A Meta-Analytic Review. *Cognitive Behavior Therapy*, 38, 66 – 82.
- Chen, D. and Manning, C. (2014). A Fast and Accurate Dependency Parser Using Neural Networks. *In: Proceedings of EMNLP*.
- Da Silva, N. F., Hruschka, E. R., & Hruschka, E. R. (2014). Tweet sentiment analysis with classifier ensembles. *Decision Support Systems*, 66, 170-179.
- Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., & Stern, A. (2004). Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *Bmj*, 328(7449), 1166.
- Johnsen, J. A. K., Rosenvinge, J. H., & Gammon, D. (2002). Online group interaction and mental health: An analysis of three online discussion forums. *Scandinavian Journal of Psychology*, 43(5), 445-449.
- Kauer, S. D., Mangan, C., & Sanci, L. (2014). Do online mental health services improve help-seeking for young people? A systematic review. *Journal of Medical Internet Research*, 16(3). doi: 10.2196/jmir.3103
- Keller, M. B. (2006). Prevalence and impact of comorbid anxiety and bipolar disorder. *The Journal of clinical psychiatry*, 67, 5-7.

- Korkontzelos, I., Nikfarjam, A., Shardlow, M., Sarker, A., Ananiadou, S., & Gonzalez, G. H. (2016). Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. *Journal of biomedical informatics*, 62, 148-158.
- Kuester A., Niemeier H., Knaevelsrud C. (2016). Internet-based interventions for posttraumatic stress: A meta-analysis of randomized controlled trials. *Clinical Psychology Review*, 43, 1-16.
- Kummervold, P. E., Gammon, D., Bergvik, S., Johnsen, J. A. K., Hasvold, T., & Rosenvinge, J. H. (2002). Social support in a wired world: use of online mental health forums in Norway. *Nordic journal of psychiatry*, 56(1), 59-65.
- Lufi & Parish-Plass, J. (2011). Sport-based rroup therapy program for boys with ADHD or with other behavioral disorders. *Child & Family Behavior Therapy*, 33(3), 217-230.
- Malmasi, S., Zampieri, M., & Dras, M. (2016). Predicting post severity in mental health forums. *Order*, 2, 8.
- Mathers, C. (2008). The global burden of disease: 2004 update. *World Health Organization*.
- Mitchell, J. T., Sweitzer, M. M., Tunno, A. M., Kollins, S. H., & McClernon, F. J. (2016). "I use weed for my ADHD": a qualitative analysis of online forum discussions on cannabis use and ADHD. *PLoS one*, 11(5).
- Motlova, L. (2007). Schizophrenia and family. *Neuro endocrinology letters*, 28, 147-159.
- Neural Networks Dependency Parser (n. d.). Stanford. Accessed Oct. 15, 2017. Retrieved from <https://nlp.stanford.edu/software/nndep.html>
- Nguyen, T., Phung, D., Dao, B., Venkatesh, S., & Berk, M. (2014). Affective and content analysis of online depression communities. *IEEE Transactions on Affective Computing*, 5(3), 217-226.
- Perich, T., Manicavasagar, V., Mitchell, P. B., & Ball, J. R. (2013). The association between meditation practice and treatment outcome in Mindfulness-based Cognitive Therapy for bipolar disorder. *Behaviour research and therapy*, 51(7), 338-343.
- Riper, H., Blankers, M., Hadiwijaya, H., Cunningham, J., Clarke, S., Wiers, R., ... & Cuijpers, P. (2014). Effectiveness of guided and unguided low-intensity internet interventions for adult alcohol misuse: a meta-analysis. *PLoS One*, 9(6).
- Roehrig, C. (2016). Mental disorders top the list of the most costly conditions in the United States: \$201 billion. *Health Affairs*, 35(6), 1130-1135.
- Salas-Zárate, M. del P., Medina-Moreira, J., Lagos-Ortiz, K., Luna-Aveiga, H., Rodríguez-García, M. Á., & Valencia-García, R. (2017). Sentiment Analysis on Tweets about Diabetes: An Aspect-Level Approach. *Computational and Mathematical Methods in Medicine*. <http://doi.org/10.1155/2017/5140631>
- Saleem, S., Pacula, M., Chasin, R., Kumar, R., Prasad, R., Crystal, M., ... & Speroff, T. (2012). Automatic detection of psychological distress indicators in online forum posts. In: *Signal & Information Processing Association Annual Summit and Conference*, 1-4.
- Solomon, O. (2010). What a dog can do: Children with autism and therapy dogs in social interaction. *Ethos*, 38(1), 143-166.
- Spijkerman, M. P. J., Pots, W. T. M., & Bohlmeijer, E. T. (2016). Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials. *Clinical psychology review*, 45, 102-114.
- Suicide data (n. d.). 2017. Retrieved from http://www.who.int/mental_health/prevention/suicide/suicideprevent/en/
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267-307.
- 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. In: *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 201-213.
- Weathers, F. W., Keane, T. M., & Davidson, J. R. (2001). Clinician-Administered PTSD Scale: A review of the first ten years of research. *Depression and anxiety*, 13(3), 132-156.
- Wilens, T. E. (2004). Impact of ADHD and its treatment on substance abuse in adults. *Journal of Clinical Psychiatry*, 65, 38-45.

Zanarini, M. C., Frankenburg, F. R., Dubo, E. D., Sickel, A. E., Trikha, A., Levin, A., & Reynolds, V. (1998). Axis I comorbidity of borderline personality disorder. *American Journal of Psychiatry*, 155(12), 1733-1739.