Introduction

• This paper focuses on online Mental Health (MH) communities and studies how users’ contributions to these communities change over one year.

• We define a metric called the Mental Health Contribution Index (MHI), which we use to measure the degree to which users’ contributions to mental health topics change over the one-year period.

• We seek to address three research questions:

  RQ1. How do users, grouped by their MHI scores, express different symptoms of MH problems throughout the year in general?

  RQ2. Can we build a classifier to predict if a user’s contributions to MH subreddits will increase or decrease during the second half of the year?

  RQ3. What factors from the first six months correlate with either an increase or a decrease in MH contributions in the second half of the year?

Data

• Aim: find three groups of users whose contributions to MH communities increase (Increase Group), decrease (Decrease Group) or stay the same (No Change Group) over time.

• Method: Crawl data through the Python Reddit API PRAW and filter target user groups through MHCI. Finally, manually rule out users without self-reported diagnoses of MH problems.

• Result: Identify 641 users for Increase Group, 758 users for Decrease Group and 368 users for No Change Group from 53,416 redditors.

\[ MHI_C(r) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\text{score}(r)}{\text{score}(i)} \right) + \frac{1}{m} \sum_{j=1}^{m} \left( \frac{\text{score}(r)}{\text{score}(j)} \right) \]

\[ MHC(r) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\text{score}(r)}{\text{score}(i)} \right) + \frac{1}{m} \sum_{j=1}^{m} \left( \frac{\text{score}(r)}{\text{score}(j)} \right) \]

\[ \text{MHI}(r) = MHC(r) \times MHI_C(r) \]

RQ1: Changes of MH Symptoms

• Emotional Symptoms: High MHCI users show increased use of negative emotion categories and decreased use of positive emotion categories.

• Linguistic Symptoms: High MHCI users show an increased use of verbs, which is considered to positively correlate with sensitivity, increased use of “1st person singular” and decreased use of “2nd person plural” and “3rd person singular pronouns” which indicate that they become more socially isolated and self-attentional.

• Subjectivity Symptoms: High MHCI users tend to increase negative opinions in the second half of the year.

RQ2: A Classification Task

• Aim: distinguish between high and low MHCI users based on only the texts that the users have written.

• Features: AverageWord2Vec, Average GloVe, Doc2Vec and LIWC.

• Classifiers: Logistic Regression (LR), Support Vector Machine (SVM) and a custom Neural Network (NN).

RQ3: Factors Correlate with MH Changes

• Algorithm 1: We use propensity score matching (PSM) to find contributions to certain subreddits in t<sub>1</sub> to correlate with increased MH (high MHCI) or decreased MH (low MHCI) contributions to MH subreddits in t<sub>1</sub>.

• Algorithm 2: We use LDA classification (LDA) to find contributions to certain subreddits in t<sub>1</sub> to correlate with increased MH (high MHCI) or decreased MH (low MHCI) contributions to MH subreddits in t<sub>1</sub>. The list contains high to low MHCI contributions to the subreddits.

Insightful Findings

• Support Communities: Support communities are shown to correlate with decreased MH contributions in t<sub>1</sub>, which is shown to be correlated with reduced MH symptoms in RQ1. MH support subreddits include /r/depression, /r/BipolarReddit, /r/SuicideWatch, /r/bipolar/, /r/mentalhealth/, /r/socialanxiety/ and /r/BPD (Borderline Personality Disorder). Other support subreddits include /r/tapecounseling (help with sexualized violence), /r/StopGaming (help with video game addiction) and /r/pornfree (help with addiction to porn).

• Interesting Pictures, Comics and Memes: Some subreddits focus on sharing images, captioned photos etc. that are intended to be funny. This category includes /r/jokes and /r/comics. We found both similar decreases in MH contributions in t<sub>1</sub>.

• Story Sharing and Friend Making: These subreddits correlate with decreased MH contributions. /r/LongDistance is a subreddit to share stories about long-distance relationships and /r/Reteme for users to rate everyone else.

• Politics: There are two subreddits related to politics in Table 4 and 5. They are /r/WeeklyPolicy and /r/AlexPolitics, and both correlate with increased MH contributions in t<sub>1</sub>.

• Other Subreddits: /r/baseball correlates with reduced MH in subreddits in t<sub>1</sub>. /r/Juice, /r/Juice/ and /r/Juice/ correlate with increased MH contributions.

Conclusions

• Our findings show that increased MH contributions correlate with increased MH linguistic symptoms while decreased MH contributions generally show the opposite trend.

• Further, we propose a framework for building classifiers to distinguish between high and low MHCIredditors and demonstrate the effectiveness of word embeddings and document embeddings in this task.

• Our work also reveals the underlying correlation between users’ engagement in discussions in different subreddits and changes in those users’ MH contributions over time.

Acknowledgements