Text-Based Detection and Understanding of Changes in Mental Health

Yaoyiran Li, Rada Mihalcea, and Steven R. Wilson, University of Michigan

{yaoyiran,mihalcea,steverw}@umich.edu

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 (MHCI), 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 MHCI 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 about 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

$$MHCI(r) = \alpha\frac{m_2^r+1}{m_1^r+1} + \beta\frac{(m_2^r+1)(n_1^r+1)}{(m_1^r+1)(n_2^r+1)}$$

$$increase\ group: MHCI(r) > 2.5, MHCI'(r) > 5$$

$$decrease\ group: MHCI(r) < 0.4, MHCI'(r) < 0.2$$

$$no\ change\ group: 0.9 < MHCI(r) < 1.1, 0.75 < MHCI'(r) < 1.25$$

r: a user; m_1, m_2: MH contributions in two half-years; n_1, n_2: NonMH contributions in two half-years

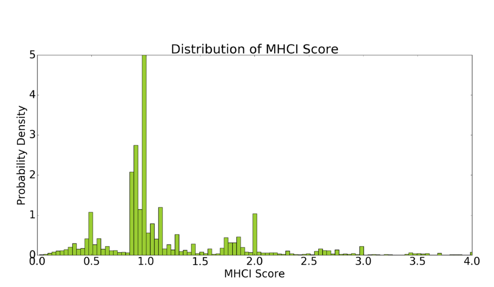

Fig. 1: Distribution of MHCI score in 53, 416 redditors

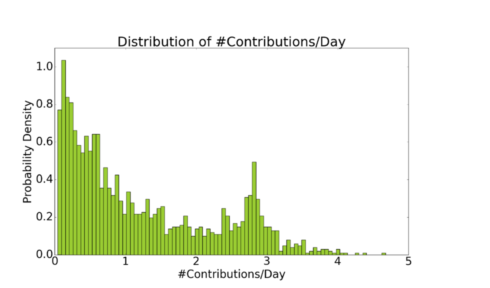

Fig. 2: Distribution of number of contributions per day over 1,767 redditors.

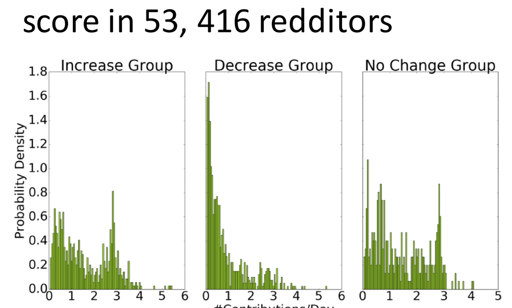

Fig. 3: Distribution of number of contributions per day over three types of users respectively.

RQ1: Changes of MH Symptoms

Category	Time Period 1	Time Period 2	t-stat	p
negate	0.0223	0.0242	21.569	***
death	0.0022	0.0024	6.420	***
health	0.0079	0.0100	39.353	***
affect	0.0633	0.0662	19.994	***
leisure	0.0128	0.0116	-18.577	***
interrogative	0.0174	0.0176	3.001	-
adverb	0.0618	0.0640	20.505	***
conjunction	0.0703	0.0721	11.791	***
pronoun	0.1685	0.1779	42.618	***
verb	0.1827	0.1901	32.261	***
1st person singular	0.0595	0.0651	39.597	***
1st person plural	0.0051	0.0047	-10.564	***
2nd person	0.0202	0.0217	17.077	***
3rd person singular	0.0131	0.0126	-7.080	***
positive emotion	0.0368	0.0365	-2.795	-
negative emotion	0.0258	0.0287	30.963	***
sad	0.0046	0.0058	28.839	***
anxiety	0.0037	0.0046	23.795	***

Table 1: Welch's t-test results on LIWC semantic categories between contents of two six-month periods for increase group users.

Category	Time Period 1	Time Period 2	t-stat	p
negate	0.0248	0.0231	-16.581	***
death	0.0027	0.0023	-9.489	***
health	0.0110	0.0081	-44.124	***
affect	0.0691	0.0623	-28.521	***
leisure	0.0103	0.0120	22.981	***
interrogative	0.0179	0.0177	-1.748	-
adverb	0.0658	0.0619	-23.448	***
conjunction	0.0735	0.0711	-13.696	***
pronoun	0.1873	0.1687	-71.382	***
verb	0.1944	0.1833	-41.384	***
1st person singular	0.0739	0.0573	-99.840	***
1st person plural	0.0047	0.0055	20.309	***
2nd person	0.0217	0.0197	-20.652	***
3rd person singular	0.0120	0.0123	3.634	**
positive emotion	0.0364	0.0366	1.747	-
negative emotion	0.0317	0.0269	-41.996	***
sad	0.0077	0.0049	-54.379	***
anxiety	0.0051	0.0039	-28.349	***

Table 2: Welch's t-test results on LIWC semantic categories between contents of two six-month periods for decrease group users.

- **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 '1st 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 these users have written.
- **Features**: AverageWord2Vec, Average GloVe, Doc2Vec and LIWC.
- **Classifiers**: Logistic Regression (LR), Support Vector Machine (SVM) and a custom Neural Network (NN)

Feature and Classifier	Accuracy	Precision	Recall	F1-Score
Word2Vec+LR	0.7713 (+/- 0.0662)	0.7542 (+/- 0.0758)	0.7442 (+/- 0.0986)	0.7485 (+/- 0.0760)
Word2Vec+SVM	0.7820 (+/- 0.0688)	0.7521 (+/- 0.0760)	0.7831 (+/- 0.0911)	0.7668 (+/- 0.0747)
Word2Vec+NN	0.7649 (+/- 0.0842)	0.7454 (+/- 0.0946)	0.7457 (+/- 0.1500)	0.7395 (+/- 0.1095)
GloVe+LR	0.7756 (+/- 0.0672)	0.7638 (+/- 0.1046)	0.7442 (+/- 0.0584)	0.7530 (+/- 0.0655)
GloVe+SVM	0.7692 (+/- 0.0586)	0.7485 (+/- 0.0822)	0.7505 (+/- 0.0658)	0.7489 (+/- 0.0608)
GloVe+NN	0.7527 (+/- 0.0747)	0.7436 (+/- 0.0907)	0.7209 (+/- 0.1167)	0.7269 (+/- 0.0877)
LIWC+LR	0.8142 (+/- 0.0412)	0.7941 (+/- 0.0605)	0.8066 (+/- 0.1327)	0.7980 (+/- 0.0574)
LIWC+SVM	0.8185 (+/- 0.0491)	0.8052 (+/- 0.0584)	0.7989 (+/- 0.1219)	0.8004 (+/- 0.0631)
LIWC+NN	0.8235 (+/- 0.0419)	**0.8170 (+/- 0.0639)**	0.8238 (+/- 0.1066)	0.8143 (+/- 0.0774)
Doc2Vec+LR	0.8234 (+/- 0.0668)	0.8107 (+/- 0.0610)	0.8019 (+/- 0.1187)	0.8054 (+/- 0.0803)
Doc2Vec+SVM	0.8113 (+/- 0.0350)	0.7955 (+/- 0.0399)	0.7925 (+/- 0.0769)	0.7934 (+/- 0.0446)
Doc2Vec+NN	**0.8241 (+/- 0.0377)**	0.8088 (+/- 0.0563)	**0.8268 (+/- 0.0691)**	**0.8181 (+/- 0.0342)**
Doc2Vec+LIWC+LR	0.8392 (+/- 0.0610)	0.8284 (+/- 0.0733)	0.8207 (+/- 0.1025)	0.8235 (+/- 0.0699)
Doc2Vec+LIWC+SVM	0.8306 (+/- 0.0666)	0.8104 (+/- 0.0689)	0.8238 (+/- 0.1060)	0.8163 (+/- 0.0758)
Doc2Vec+LIWC+NN	**0.8614 (+/- 0.0535)**	**0.8587 (+/- 0.0544)**	**0.8519 (+/- 0.0763)**	**0.8558 (+/- 0.0333)**

Table 3: Classification results with 10-fold cross-validation. We report here the average accuracy, precision, recall, f1-score and their 95% confidence interval of the score estimate (i.e. 2 times standard deviation).

RQ3: Factors Correlate with Changes in MH Contributions

Algorithm 1 Measuring the effects of treatments on MHCI

Require: $X_i, Y_i, T_{j,i}, i = 1$ to n, $j = 1$ to m
  for $j = 1$ to m **do**
    step 1: Split data into a training set and a test set.
    step 2: Fit $model_j$ to the training data.
    step 3: Form treatment group and control group in test set based on Propensity Score Matching.
    step 4: Conduct Welch's t-test on treatment and control groups.
  end for
  return t-stats for all m treatments

Algorithm 1: We use propensity score matching (PSM) to find if contributions to certain subreddits in t_1 correlate with increased (high MHCI) or decreased (low MHCI) contributions to MH subreddits in t_2. X_i's are confounding variables, $T_{j,i}$'s are treatment labels and Y_i's are user labels.

Treatment	t-stat
r/WikiLeaks	3.464
r/vancouver	3.464
r/trypophobia	2.752
r/Marijuana	2.738
r/Ask_Politics	2.449
r/cordcutters	2.449
r/piercing	2.291
r/cars	2.254
r/announcements	2.190
r/MeanJokes	2.038
r/AskUK	2.070
r/Bandnames	2.000
r/solotravel	2.000
r/whatisthisthing	1.981
r/Bitcoin	1.951

Table 4: Top 15 treatments that correlate with an increase in MH contributions.

Treatment	t-stat
r/depression	14.191
r/BipolarReddit	5.740
r/SuicideWatch	5.554
r/StopGaming	4.472
r/bipolar	4.354
r/pics	4.157
r/mentalhealth	4.057
r/pornfree	3.464
r/rapecounseling	3.314
r/baseball	3.162
r/socialanxiety	3.004
r/comics	2.758
r/LongDistance	2.738
r/Rateme	2.662
r/BPD	2.660

Table 5: Top 15 treatments that correlate with a decrease in MH contributions.

Insightful Findings

- **Support Communities:** Support communities are shown to correlate with decreased MH contributions in t_2 which are 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 communities include 'r/rapecounseling' (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/pics' and 'r/comics' and both correlate with decreased MH contributions in t_2.
- **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/Rateme' for users to rate everyone else.
- **Politics**: There are two subreddits related to politics in Table 4 and 5. They are 'r/WikiLeaks' and 'r/Ask Politics', and both correlate with increased MH contributions in t_2.
- **Other Subreddits** : 'r/baseball' correlates with reduced MH contributions in t_2. 'r/Marijuana', 'r/trypophobia' (a community for those with a common fear of irregular clusters of holes or bumps found in the world) and 'r/piercing' (for discussion of
- various body piercings and jewelry) 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 MHCI redditors 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

We thank all anonymous reviewers for their constructive suggestions on our work. We also thank Dr. M´arcio Duarte Albasini Mour˜ao for helpful discussions with us on RQ1. This work is partly supported by the Michigan Institute for Data Science, by the National Science Foundation under grant #1344257 and by the John Templeton Foundation under grant #48503.