Predicting Mental Health Status on Social Media

A Preliminary Study on Microblog

Bibo Hao, Lin Li, Ang Li, Tingshao Zhu[1]

{Institute of Psychology , University of Chinese Academy of Sciences}, CAS
{haobibo12,lilin510,liang08}@mails.ucas.ac.cn
tszhu@psych.ac.cn

Abstract. The rapid development of social media brings about vast user generated content. Computational cyber-psychology, an interdisciplinary subject area, employs machine learning approaches to explore underlying psychological patterns. Our research aims at identifying users' mental health status through their social media behavior. We collected both users' social media data and mental health data from the most popular Chinses microblog service provider, Sina Weibo. By extracting linguistic and behavior features, and applying machine learning algorithms, we made preliminary exploration to identify users' mental health status automaticly, which previously is mainly measured by well-designed psychological questionnaire. Our classification model achieves the accuracy of 72%, and the continous predicting model achieved correlation of 0.3 with questionnaire based score.

Keywords: microblog, mental health, prediction, automation

1 Introduction

The emergence of social media has been a highlight in the rapid development of the Internet, which attracts many users to express themselves through Social Media like Facebook, Twitter and so on. Users' online behaviors are fabricating a cyber-space, which reflects and interacts with the real world. Hence, users' online behavior can be an indicator of their psychological characteristics in real life.

Nowadays, more and more people are suffering from mental disorders like depression, anxiety, tension, etc., due to pressures, external environment and other reasons. These mental disorders may influence users' life severely and sometimes even may lead to suicides.

In the past, people with mental health problems may be advised to consult a therapist, or they may look for psychotherapy initiatively. People realize mental health disorders by intuition or self-reported questionnaires. Even if mental health problem is realized, psychotherapy may be unreachable due to lack of resources.

[*] Corresponding Author.

HCI, p. 1,2013.

Nowadays, cyber-space is providing a new approach to change such situation. Since worldwide used SNSs have been a part of users' life, we propose to identify users' mental health problem through their online behavior and expression, and give them appropriate suggestions when necessary.

The rest of this paper is organized as follows. In Section 2, we introduce some related work. Section 3 will describe the experiment procedure and dataset in detail. Section 4 will talk about our modeling thought. In Section 5, we present the detailed algorithm. Experiment results will be discussed in Section 6. At last, we will give a conclusion in Section 7.

2 Related Work

The research of mental health has been a relatively ripe area in psychology, while recent researches have made exploration to migrate conclusions to cyber-space.

Personality, a stable psychological trait, is a fundamental research area in psychology. Therefore, early researches in cyber-psychology tried to identify and predict users' personality online. François et al. discovered that users' personality can be predicted through their linguistic cues [3], whose work is based on the corpus of essay and EAR. Jennifer [4], Yoram [5]. et al. tried to predict users' personality based on Facebook, took user's behavior into consideration. Their work is based on the assumption that users present themselves on Facebook profile in the way in accordance with their personality, while such assumption may not be applicable to other non-real-name SNSs. Daniele et, al have studied to predict personality with Twitter [6].These works has proved that it is feasible to predict psychology trait through social media.

The difference between personality and mental health status is that, personality is a relatively stable trait while mental health status is much easier to change than personality. One's personality is shaped when he is grown up and doesn't change too much during his or her life, whereas one's mental health changes along with his or her mind, external environment and other reasons.

Fan et al. made a survey on the relationship between web behaviors and mental health [7].It has been proved that, web behaviors like online time, instant messenger usage, browsing particular content, activity frequency etc., are closely related to mental health status. Dong et al. use web browser to predict users' mental health status and recommend content to alleviate mental disorders [7].Their application proves the feasibility to predict mental health status through online behaviors. Yet it is notable that their approach record users' browsing data, which may leads to privacy concerns.

Michael et al. used "Weibo use intensity and motivation" as behavior factors to analysis the political efficacy among young Chinese citizens [9]. Their work reveals the relationship between Weibo behaviors and users' opinions, while the behavior features are measured in conventional questionnaire.

Our research goal is to find a way to predict users' mental health status through users' publicly accessible data or behaviors.

3 Experiment Design

Sina Weibo is a leading microblog service provider in China with more than 300 million users, and Weibo users post more than 100 million microblog publicly every day [1]. Therefore, we conduct our study on Sina Weibo, as the platform to collect data. On Sina Weibo, people can pick their nick names instead of real name. Since the microblogs are in public, it is accessible to all Weibo users or crawler. However, each microblog is less than 140 characters, which makes it challengeable to analysis using conventional methods.

Predicting users' mental health status can be regarded as a typical machine learning problem. In this problem, feature space includes features extracted from users' Weibo behavior data, and the prediction variable is users' mental health statuses. To carry out the prediction task, we need to train a classifier based on labeled data, and apply the classifier on unlabeled data to predict the mental health status. In our case, labeled data refer to users' Weibo data labeled with the correspondent users' mental health statuses.

Users' Weibo data can be accessed publicly, since Weibo provides APIs to download users' data, including personal profile, microblogs, followings, followers and so on. We have implemented a Weibo crawler, which can download specified users' data conveniently. Whereas users' mental health statuses are costly to obtain. We use SCL-90-R, an influential and widely used psychometric instrument to assess users' mental health status [10], which consists of 90 questions. When a subject fills the questionnaire, it yields 9 scores correspond to 9 mental health dimensions: anxiety, depression, somatization, obsessive-compulsive, interpersonal sensitivity, hostility, photic anxiety, paranoid ideation and psychoticism. We have published a Weibo app, *PsyMap* (http://ccpl.psych.ac.cn:10002), for users to fill the questionnaire and collect filling data.

When a Weibo user agrees to participate our experiment, he or she logins into our Weibo app by an OAuth2.0 based authorization interface provided by Sina Weibo. The user is instructed to complete our online SCL-90-R questionnaire. At the same time, Weibo crawler will download his or her Weibo data, ready for features extraction.

During modeling procedure, feature vectors, extracted from users' Weibo data, are dependent variables, and users' SCL-90-R assessments are independent variables. Our mental health predicting model is trained based on these data. The whole procedure is illustrated in Fig. 1.

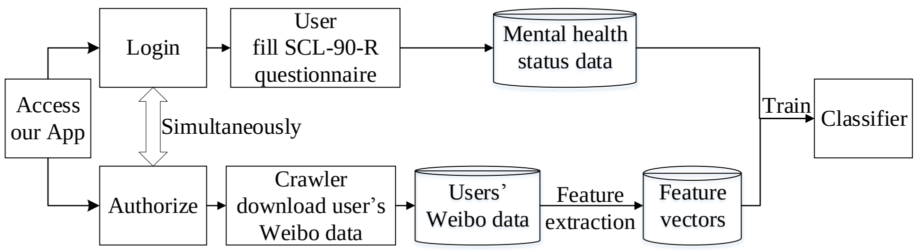


Fig. 1. Experiment Procedure

In June and July of 2012, we recruited 448 Weibo volunteers to take part in our study, and the recruitment is conducted randomly in order to get a normal user distribution. As a preprocess procedure, we scale 448 subjects' SCL-90-R scores by normalization $S_{scaled} = \frac{S_{original} - avg_{original}}{std_{original}}$, i.e. the scaled score S_{scaled}, ranges roughly from -1 to 4, has the average value of 0 and the standard deviation of 1. Higher score corresponds to poor mental health status. Fig. 2 shows the distribution of scaled 9 mental health dimension score of our dataset. We can see that most users' scaled scores are less than 1, in other words, most users have a good mental health status, while users scored greater than 1 tend to have poorer mental health status. In our dataset, users scored greater than 1 amounts for 15% roughly.

Additionally, SCL-90-R yields scores of 9 dimensions, so we build predicting model for each dimension separately following the same procedure, only the predicting variable needs to be changed in the training process.

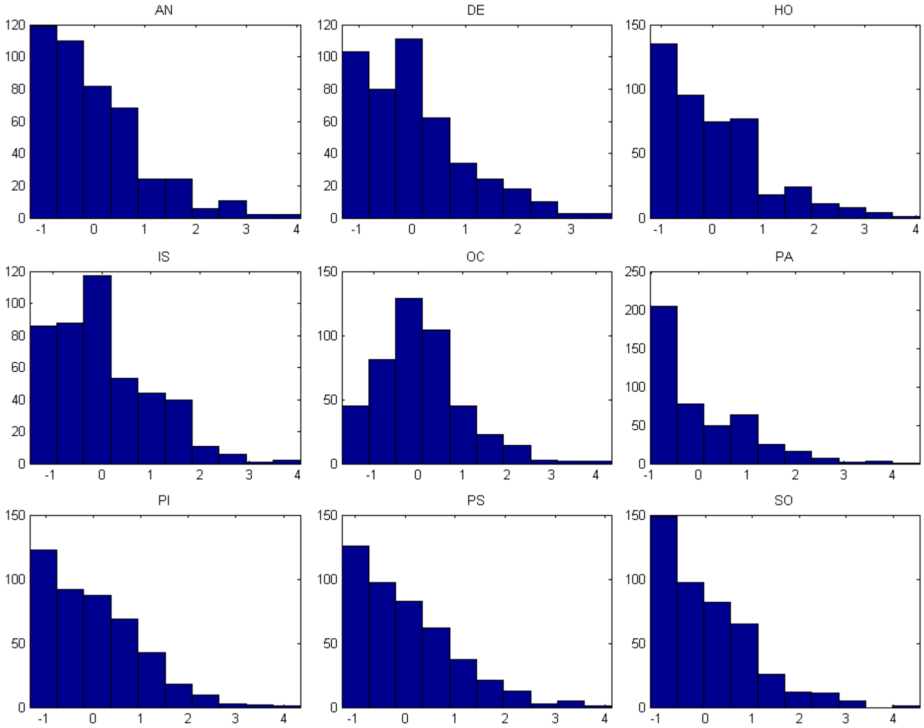


Fig. 2. Distribution of 9 mental health dimension score of our dataset

4 Modeling

4.1 Feature Extraction

Features are actually even more important than algorithms in modeling procedures. Obviously, features relevant to users' mental health status may leads to better prediction results. We extract features from two perspectives, the first is users' microblog content, and the second is user's profile.

Generally, feature extraction needs particular domain knowledge from a different extent. In our study, we made two explorations to extract features. Firstly, we adopted little domain knowledge and establish model 1, which can be regarded as a "bootstrap" method. In this model, features are words that may be correspondent to mental health disorders. These words are discovered from the corpus by a vote whose core idea is based on Naïve Bayes principle. Secondly, model 2 introduced word dictionary, SNS profile and behavior features as domain knowledge. Different from model 1, in this model we are employing some psychological conclusions. Previous researches have proved that psychological traits or statuses are connected to linguistic expression and online behavior, so we extract features artificially in model 2. We also compare these two models at last.

4.2 Model Selection

Various effective machine learning algorithms have been developed in last decades, among which, SVM and neural network (NN) are thought to be best ones, yet it is worth to notice that, in practice they are both non-linear models. However, in psychological analysis, when a model is involved with multiple dependent variables, a non-linear model will be difficult to interpret, let alone complicated models like SVM and NN. The purpose of psychological analysis is to reveal the interpretable relationships between behaviors and psychology traits or status. Hence, SVM and NN are like black box, even if their prediction are nearly perfect, it will make little sense to understand the relationship between users' behavior features and mental health status.

As an exploring experiment, we choose to employ Naïve Bayes and linear machine learning algorithms, in order to generate models can not only predict in a fair accuracy but also easy to interpret. Naïve Bayes models are intuition-driven, and linear models can reveals the correlation between variables explicitly.

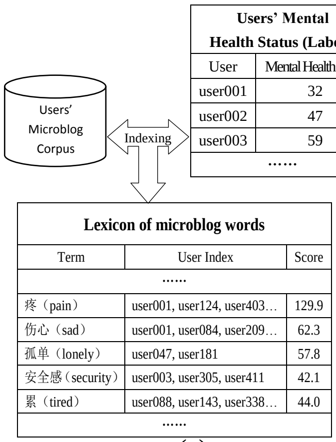
5 Algorithms and Prediction Models

In this part, we put our modeling ideas into practice. Model 1 employs the "bootstrap" feature extraction method and Naïve Bayes principle. Model 2 extracts linguistic features based on well-built dictionary, and extracts behavior features from user's profile. Model 2 uses linear machine learning algorithm, namely linear regression and pace regression.

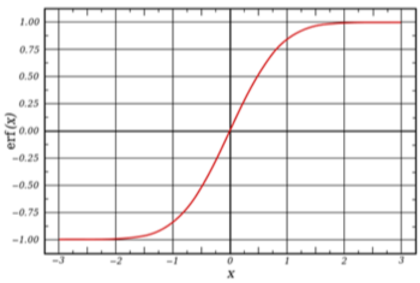
5.1 Model 1 – Predictor based on Naïve Bayes Predisposition Words Selector

Weibo users express their views, feelings, mood and so on by words, expression images, photos, etc. Intuitively, we believe that particular terms may link to particular mental health status. Hence, we make our model to vote for predisposition words that low mental health status users use more, and apply these words to predict low mental health status.

Fig. 3 shows the procedure of building lexicon with mental health weights. As we mentioned, each user is labeled with a mental health status score, so we use this score to label the words used by this user.



(a)



$$\text{Score(term)} = \sum_{i=1}^{N_{Users}} \frac{1}{1 + \exp(-r * U_i * n_{i,term})}$$

(b)

Fig. 3. Building Lexicon with Mental Health Weights

Firstly, for each user i with the mental health score of U_i, we pick up all this user's microblogs, and for each unique term in these microblogs, we label the term with a mental health score U_i and also record the term frequency $n_{i,term}$, in other words, user i use the term for $n_{i,term}$ times. So each term has several pairs of $(U_i, n_{i,term})$.

Secondly, we apply the formula in Fig. 3(b) to figure out the mental health score of each terms in the lexicon. Here, our formula use a sigmoid function, in order to eliminate affections caused by the situation where a term is used too many times by some users. The formula contains a parameter r, which adjust the steepness of sigmoid function. To a large r, terms score goes saturated when term frequency is not very high, whereas to a small r, saturation occurs only when term frequency is very high. In our case, we choose $r = 0.4$ as an empirical value. When applying the formula, we also set up a condition, that is only those terms used by more than 3 users are put in the lexicon. By doing so, we want the terms to have sufficient support.

It can be seen that, when a term is more often used by poor mental health status users, it will score high, when a term is more often used by good mental health status users, it will score negative, and when a term is often used by both good and poor mental health status users, its score will be eliminated to approach 0.

疼(pain)	破(bad)	对不起(sorry)	纠结(dilemma)	可是(but)	命(fate)	
孤单(lonely)	久(long time)	贱(mean)	困(tired)	心里(in one's mind)		
只是(but)	伤心(sad)	心疼(sad)	鼻(nose)	面(face)	难受(sad)	修(fix)
讨厌(boring)	失眠(insomnia)	不好(not good)	头疼(headache)			
安全感(security)	闹心(worrying)	人家(a special form of first person pronoun)				

Fig. 4. Some top-ranked terms in the depression lexicon

After the lexicon with mental health weights is built, we can use it to predict unlabeled users' mental health status. Since the above procedure is based on the Naïve Bayes principle, we conjecture terms in the lexicon with high weights should indicate metal health disorders. This is supported by the experiment results, as we can see in Fig. 4, the listed terms are some top ranked in the depression lexicon. Their meanings are closely linked to poor mental health status by semantics and common sense.

In the prediction procedure, we adopted top 1000 out of more than 300 thousand terms, in the lexicon, since they are more closely related to mental health disorders. For an unlabeled user, we just pick up all his or her microblogs, and for all terms he used, we search it in our selected lexicon. To those terms occurred in our lexicon, we just add up their weights to get a total value as the prediction score P.

The prediction score P needs a threshold P_T to identify whether this users is suffering from mental health disorders. In psychometrics, subjects scoring higher than $avg + std$ (in our case, the value is 1) are tend to have mental health problems. To set the value of P_T, we use the prediction procedure to "predict" users that are already labeled, i.e. the training data. We choose 50 users whose labeled scores are closest to 1, and "predict" their mental health score using the above procedure. Thus we get their prediction scores $\{P_1, P_2, \cdots, \ P_{50}\}$, then we set P_T to be the average of these scores.

At last, given a Weibo user's data, our prediction model yields user's prediction score P. When $P > P_T$, we say the odds this user has mental health problem is fairly high. This prediction model concludes as a two-class classifier. Our model implementation uses ICTCLAS [11] for Chinese lexical analysis and turn a microblog into terms.

5.2 Model 2 – Predictor Using Linguistic and Behavior Features

Model 1 is "bootstrap" since the features extraction is completed by discovering predisposition words from the corpus automatically, with little domain knowledge. In model 2, we are adopting domain knowledge and extracting features artificially.

Former research has introduced dictionary to predict psychology traits like personality. This is based on the idea that words occurred in the content is related to psychological traits or statuses. LIWC [12] is a widely used toolkit in psycholinguistics, which contains a dictionary that categorize words by psychological experiences. In our experiment, we use the simplified Chinese version of LIWC dictionary, translated from traditional Chinese version developed by Chin-Lan [13]. To assess users' emotional status, we also introduced a Chinese emotion word ontology dictionary, constructed by Linhong et al [14]. When applying dictionaries in feature extraction, we simply count the word frequency of each category occurred in user's microblog as the feature vector.

LIWC Features	Emotion word ontology Features	User profile and behavior Features
Swear Interjection Body Money Home … (78 in all)	NA: 愤怒(Angry) NB: 悲伤(Sad) NJ: 失望(Disappointed) NH: 疚(Regretful) NI: 慌(Flurried) … (21 in all)	Count of Friends Count of Followers Count of Bi-Followers Count of Microblogs Original Microblog Ratio … (38 in all)

Table 1. Some features used in model 2

Apart from linguistic features, behavior features are also important because user's behavior patterns change along with his or her mental health status.

Then we are selecting from these 137 features, and applying linear machine learning algorithms using selected ones as independent variables, to predict the mental health status score. To select features, we applied stepwise regression algorithm to choose predictable variables. Then we use the selected features to fit the mental health score using linear regression and pace regression algorithm. In stepwise regression, the parameters of penter and premove are set to 0.05 and 0.10 respectively.

6 Experiment Results

To model 1, building the lexicon for each mental health dimension is time consuming. So in our experiment we only tried on the dimension of depression. The classification accuracy on training data is 72%.

In model 2, for 9 mental health dimensions, 4 to 12 variables are selected when using stepwise regression, and the results are listed in Table 2. In the table, RAE refers to relative absolute error, RRSE refers to root relative squared error, and Corr. refers to Pearson correlation coefficient

Dimension	Linear Regression				Pace Regression		
	RAE%	RRSE%	Corr.	R^2	RAE%	RRSE%	Corr.
AN.	97.28	95.49	0.3012	0.13	95.70	94.03	0.3344
HO.	95.78	95.69	0.2883	0.11	94.39	94.41	0.3214
O.C.	97.77	96.40	0.271	0.11	95.45	94.57	0.319
I.S.	96.01	97.29	0.2504	0.11	95.81	98.85	0.2234
PS.	97.77	97.42	0.229	0.09	98.47	99.42	0.1793
SO.	96.91	97.99	0.2148	0.09	96.78	99.37	0.1823
DE.	99.00	97.99	0.2047	0.07	99.69	99.33	0.1613
P.I.	98.10	97.86	0.1926	0.05	98.00	97.75	0.1967
PA.	99.74	101.18	0.156	0.06	98.58	99.03	0.193

Table 2. Eveluation of regression results

The current experiment results shows that our model performs not well enough. In psychological analysis, if two variable has a Pearson correlation from 0.3 to 0.4, it implies that these two variables are of low correlation. Here, in the dimensions of anxiety, hostility and obsessive-compulsive, the correlation between our prediction score and questionnaire based score have reach a weak correlation by using linear regression or pace regression. The evaluation is conducted by using 10-fold cross validation.

When applying Pearson correlation coefficient as the evaluation standard, we see that in 5 dimensions pace regression perform better than linear regression. This is because pace regression algorithm has introduced clustering analysis to evaluating the contribution of each variable. But this doesn’t guarantee a better result.

7 Conclusion

Our study in this paper aims at predicting users’ mental health status based on microblog. For this purpose, we built two models. Model 1 use little domain knowledge and can select predisposition words related to mental health disorders. The core idea of this model is a Naïve Bayes classifer, users classified as a positive class are identifeid as in poor mental health status. Model 2 introduced domain knowledge, as a typical machine learning problem, we extract features from users’ microblog data, took both lingustic and behavior information into consideration, and then employ linear machine learning algorithm to the extracted features and predicting value. Model 1 is not good at contious value prediciton, while model 2 can turn into a two-class classifier simply by setting a thershold. Model 1 can also yield a lexicon with mental health status releated weights, this can be generalized to other areas where terms are connceted to predicting variable.

Experiment results shows that these models can basically predict some mental health demisions such as anxiety. This paper presents our innovation in conventional psychology research, that is to take advantage of vast social media data and machine learning methods, to predict pyschology trait or status, which can be a supplement of conventional psychometric tools - questionnaires.

As a preliminary exploration, major improvements are still to be made in our work. First, mental health status is time-varying, our models in this paper didn’t take the time factor into condiseration and use all users’ data since registeration for prediction. This might be the reason why our expirment results are not good enough. Second, our feature extracion methods are heuristic and focous on users’ static behaviors like profile, all microblogs posted by users. Time related behavior pattern features are not included in our model. Since mental health status affect behavior patterns a lot, such features should be introduced in feature works. Online microblog behaviors should also be defined more elaborately, which can be a guideline of extracting behavioral features.

8 Acknowledgments

The authors gratefully acknowledges the generous support from National High Technology Research and Development Program of China (2013AA01A606), NSFC (61070115), Institute of Psychology (113000C037), Strategic Priority Research Program (XDA06030800) and 100-Talent Project (Y2CX093006) from Chinese Academy of Sciences.

References

1. Meltzer, H., Rebecca, G., Robert, G., Tamsin, F.: Mental health of children and adolescents in Great Britain. International Review of Psychiatry, 15(1-2):185–187, 2003.
2. Xinhua News: Sina Weibo's registered user reached 300 millions. http://news.xinhuanet.com/tech/2012-02/29/c_122769084.htm (2012)
3. François, M., Marilyn, W., Matthias, M., Roger, M.: Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. Journal of Artificial Intelligence Research 30, 457-500 (2007)
4. Jennifer, G., Cristina, R., Karen, T.: Predicting Personality with Social Media. Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems. ACM, 2011.
5. Yoram, B., Michal, K., Thore, G., Pushmeet, K., David, S.: Personality and Patterns of Facebook Usage. WebSci 2012 (2012)
6. Daniele, Q., Michal, K., David, S., Jon, C.: Predicting Personality with Twitter. 2011 IEEE International Conference on Privacy, security, risk and trust and IEEE International Conference on Social Computing. (2011)
7. Fan, Z., Tingshao, Z., Ang, L., Yilin, L., Xingguo, X.: A survey of web behavior and mental health. 6th International Conference on Pervasive Computing and Applications (ICPCA), 2011 (pp. 189-195). IEEE.(2011)
8. Dong, N., Yue, N., Tingshao, Z.: Predicting Mental Health Status in the Context of Web Browsing. Web Intelligence 2012 (2012)
9. Michael, C., Xuan, W., Yinqi, H., Rui, X., Tian, J.: Microblogging, Online Expression, and Political Efficacy Among Young Chinese Citizens: The Moderating Role of Information and Entertainment Needs in the Use of Weibo. Cyberpsychology, Behavior, and Social Networking. 15.7 (2012): 345-349.
10. Derogatis, L.R. & Savitz, and K.L. The SCL-90-R and the Brief Symptom Inventory (BSI) in Primary Care In. Lawrence Erlbaum Associates, 2000.
11. Huaping, Z.: ICTCLAS, Chinese Lexical Analysis System. http://ictclas.nlpir.org (2013)
12. Pennebaker, J.W., Cindy, K.C., Molly, I., Amy, G., Roger, J.B. The Development and Psychometric Properties of LIWC 2007. (2007)
13. Chin-Lan, H., Cindy, K.C., Natalie, H., Yi-Cheng, L., Yi-Tai, S., Ben, C.P.L., Wei-Chuan, C., Michael, H.B., Pennebaker, J.W.: The Development of the Chinese Linguistic Inquiry and Word Count Dictionary. Chinese Journal of Psychology 54, Volume 2, 185-201. (2012)
14. Linhong, X., Hongfei, L., Yu, P., Hui, R., Jianmei, C.: Constructing the Affective Lexicon Ontology. Journal of the China Society for Scientific and Technical Information 27(2). (2008)