17. Anatomization and Perception of Mental Disorder Because Usage of Online Social Network Data

Aaradhana Arvind Deshmukh Phd student,Aarhus University, Denmark
aadeshmukh@sinhgad.edu
Albena Mihovska, Department of Business Science and Technology, Aarhus University, Denmark
amihovska@btech.au.dk
Ramjee Prasad, Department of Business Science and Technology, Aarhus University, Denmark
ramjee@btech.au.dk

ABSTRACT

There is an excessive growth in the usage of social networking sites. Because of this an increasing number of social network mental disorders (SNMDs), such as Cyber-Relationship Addiction, Information Overload, and Net Compulsion, have been recently reported. Passively observed symptoms of these mental disorders lead to a delayed clinical intervention. In this paper, the authors argue that mining online social data provides an opportunity to find out SNMDs at an early stage. The mental state cannot be observed directly from online social data because of that it is difficult to identify SNMDs. Our approach, instead of relying on the self-revealing of those mental factors via questionnaires proposes a new system, namely, the Social Network Mental Disorder Detection (SNMDD). Our system is evaluated through a user study with number of users of the network. We perform a feature analysis and also apply SNMD in large- scale data sets and analyze the characteristics of the three types of mental disorder. We will be providing text area and a questionnaire to calculate the depression level. On the basis of depression level we will be providing basic solution which includes books, videos, music, foods, exercises, and also chatbot and messenger to overcome some easy situations.

Index Terms—*social network, mental disorder detection, feature extraction, Decision Tree classifier.*

INTRODUCTION

Mental disorder is becoming a threat to people’s health now a days. With the rapid pace of life, more and more people are feeling mentally disturbed. It is not easy to detect the user’s mental disorder early enough[1,4] .With the penetration of web-based social networking, individuals are used to share their day by day activities and interact with friends, making it possible to use online social network data for mental disorder detection[2]. In our system, we find that the user’s disorder state is closely related to that of his/her friends in social media, and we employ a large-scale dataset from real-world social platforms to systematically study the correlation of user’s disorder states and social interactions. We first define a set of mental disorder-related textual, visual, and social attributes from various aspects. Though mental disorder itself is non-clinical and common in our life, excessive and chronic disorder can be rather harmful to the people’s physical and mental health[3]. The user’s social interactions on social networks contain useful cues for stress detection.

Social psychological studies have made two interesting observations. The first is mood contagion: a bad mood can be transferred from one person to another during social interaction. The second social interaction: people are - known to the social interaction of the user. With the advancement of social networks like Twitter, Facebook and Sina Weibo, an ever increasing number of people share their every day events and moods, and interact with friends. Due to the influence of - Facebook post content attributes and social interactions become enhance factors for mental disorder detection[5],[6].

Clinical depression is a serious condition that negatively affects how a person thinks, feels, and behaves. In contrast to normal sadness, clinical depression is persistent, often interferes with a person's ability to experience or anticipate pleasure, and significantly interferes with functioning in daily life. Untreated, symptoms can last for weeks, months, or years; and if inadequately treated, depression can lead to significant impairment, other health-related issues, and in rare cases, suicide.

After detecting a disorder level, the system can recommend hospitals to the user for medical advice, with reference to a Google map and how to take precaution for avoid disorder.

MOTIVATION

These days, many people spend a large portion of time on the Internet, whether it is for work, social interaction, gathering information, or entertainment. There is a fine line between what can be considered a healthy amount of Internet usage, and what is kn own as addiction disorder [10],[11]. Disorders may contain online relationship addiction, net compulsion, information overload. Because of online social relationship there may be issues like physical absence of the person, identity issue, people may get addicted to a virtual relationship with wrong person. Likewise, the compulsive usage of social networking sites may lead to problem of addicted behavior towards net-gaming[11]. People who are net-gaming addicts cannot stop playing games (e.g. PUBG players) on the Internet.

Information overload addicts are addicted to the unlimited information that can be found on the Internet. They spend countless hours reading and organizing data they find, often times developing obsessive-compulsive tendencies. Information overload addicts will work less productively in their careers and especially in their personal lives. So, to overcome all these problems there is a need of automatic detection of these mental disorders.

Today, the identification of potential mental disorders often falls on the shoulders of supervisors (such as teachers, employers, or parents) who can observe the aforementioned symptoms better than others but only passively. There are very few notable physical risk factors, the patients usually do not actively seek medical or psychological services to reduce these symptoms. Passive observation of symptoms of an increasing number of social network mental disorders (SNMDs) is resulting in delayed clinical intervention. It is desirable to have the ability to actively detect potential SNMD users on OSNs at an early stage.

STATE OF ART

K. Rameshwaraiah, et al. [1] presented a system for detecting the users' psychological stress states from the users' weekly social media data, leveraging tweets' content as well as users' social interactions. To fully leverage both constant and social interaction information of users' tweets, we proposed a hybrid model which combines the factor graph model (FGM) with a convolution neural network (CNN).

Experimental results show that the proposed model can improve the detection performance by 6-9 percent in F1-score. By further analyzing the social interaction data and, it is possible to discover several intriguing phenomena.

Huijie Lin et al. [2] showed that long –term stress may lead to many severe physical and mental problems. Traditional psychological stress detection usually relies on the active individual participation, which makes the detection labor-consuming, time-costing and hysteretic.

A novel method is proposed for detecting psychological stress from micro blog utilizing cross-media micro blog data; Three-level system is construct to formulate the problem, and propose a middle-level representation according to psychological and art theories, which can narrow the gap between low-level cross-media features and high-level stress semantics; here designed a Deep Neural Sparse Network based classification model to solve the problem of sparse in cross-media data.

Author's investigate the social correlations in psychological stress to further improve the detection performance. Jennifer Golbeck, et al. [3] Users' Big Five personality traits can be predicted from the public information they share on Twitter. Our subjects completed a personality test and through the Twitter API, we can collect publicly accessible information from their profiles. After processing this data, we found many small correlations in the data. Using the profile data as a feature set, we were able to train two machine learning algorithms - ZeroR and Gaussian Processes to predict scores on each of the five personality traits to within 11%

- 18% of their actual value. With the ability to guess a user's personality traits, many opportunities are opened for personalizing interfaces and information.

After surveying this paper questions about how to present trusted, socially-relevant, and well-presented information to users. Andrey Bogomolov, et al. [4] The goal of this paper was to investigate the automatic recognition of people's daily stress from three different sets of data: a) People activity, as detected through their smart phones;b) Weather conditions; and c) personality traits. The problem was modeled as a 2-way classification one. The results convincingly suggest that all the three types of data are necessary for attaining a reasonable predictive power. As long as one of t hose information sources is dropped, performances drop below those of the baselines.

Automatic stress detection based on mobile phone data can take advantage of the extensive usage and diffusion of these devices, it can be applied in several real world situations and it can be exploited for a variety of applications that are delivered by means of the same device.

Sho Tsugawa, et al. [5], Depression has become recognized as a major public health problem around the world. Aim was to establish a method by which to recognize depression by analyzing the large-scale records of users' activities in social media. The extensive evaluation of effectiveness of a user's social media activities is used for estimating degree of depression. The degree of depression of Twitter users can be measured using the results of a web-based questionnaire. Several features are extracted from the activity histories of Twitter users. By using these features, we construct models for estimating the presence of active depression. Features obtained from user activities can be used to predict depression of users.

Comparing with [1-4] above references learning robust uniform features for cross- media social data by using cross auto encoders take a more time. Feng-Tso Sun, et al. [6]; A multimodal approach is used to model the mental stress activation affected by physical activities using accelerometers, ECG, and GSR sensors. Accelerometer data is necessary to improve mental stress detection in a mobile environment. Here Decision Tree classifier has the best performance using 10-fold cross validation. This activity-aware scheme for mental stress detection can facilitate the development of many affective mobile applications using physiological signals (e.g., stress management, affective tutoring, and emotion-aware human computer interfaces).

Repeated information in relevant answers requires the user to browse through a huge number of answers in order to actually obtain information. Liqiang Nie, et al. [7]; A medical terminology is used to present an assignment scheme to bridge the vocabulary gap between health seekers and healthcare knowledge.

A novel scheme is used to code the medical records by jointly utilizing local mining and global learning approaches. However, the local mining approach may suffer from information and low precision, which motivates us to propose a global learning approach to compensate for the insufficiency of local coding approach. Extensive evaluations on a real world dataset demonstrate that this scheme is able to produce promising performance as compared to the prevailing coding methods. In [7] problem faces about investigation how to flexibly organize the unstructured medical content into user needs-aware ontology by leveraging the recommended medical terminologies.

Kimberly s. Young [8], some people get addicted the internet in same way to drugs or alcohol, online social users get addicted to. Research among sociologists, psychologists, or psychiatrists has not formally identified addictive use of the Internet as a problematic behavior. this study developed a brief eight-item questionnaire referred to as a Diagnostic Questionnaire (DQ), which modified criteria for pathological gambling to provide a screening instrument for classification of participants.

Gap found in briefing of techniques to better incorporate multi-word terms and out-of-vocabulary words; advanced Natural Language Processing (NLP) techniques for learning word relations from free-form text; evaluation of latent concept relation suggestion, and predicting the type of relations. Yuan Zhang, et al. [9] focused on human emotion is one important part, which is affected by the dynamics of social networks. They proposed a system that consists of MoodCast method based on a dynamic continuous factor graph model for modeling and predicting users' emotions in a social network. Based on the information history MoodCast learns a discriminative model for predicting users' emotion status at time t. For model learning, it uses a Metropolis-

Hastings algorithm to obtain an approximate solution. Experimental results on two different real social networks demonstrate that the proposed approach can effectively model each user's emotion status and the prediction performance is better than several baseline methods for emotion prediction.

Young Min Baek, et al. [10]; tested the effect of Social Networking Sites (SNS) on the users' psychological condition, in terms of subjective loneliness, interpersonal trust, and SNS addiction. This study distinguishes types of SNS relationships, and investigates their relationships with social isolation, interpersonal trust, and SNS addiction. A gap was found in parallelization, which is considered in our algorithm in order to scale it up further.

Katherine Chak, et al. [11] found that some patterns of internet use are associated with loneliness, shyness, anxiety, depression, and self-consciousness. Here, the study attempted to examine the potential influences of personality variables, such as shyness and lack of control, online experiences, and demographics on internet addiction.

Results indicated that the higher the tendency of one being addicted to the Internet, the shyer the person is, the less faith the person has, the firmer belief the person holds in the irresistible power of others, and the higher trust the person places on the chance in determining his or her own course of life.

Kiwon Kim et al. [12], defined Internet Addiction (IA) as a psychological dependence on the internet, regardless of the type of activities once logged on. The aim of this study was to investigate the association between suicide attempts and sleep among community-dwelling adults with IA. Among adults with IA, poor sleep quality was found to be associated with more severe IA and lifetime suicide attempt. A variety of poor sleep quality indexes associated with IA were shown in [12] and the association showed a positive correlation with IA severity.

The disadvantages of this proposed system is that, this is a cross sectional study, using questionnaires and interview depending on subjective memory therefore; there are limitations to objective measurement related to sleep. This is overcome in our proposal. Based on the assumption that Net-generation has unique characteristics, the study in [13], examined (1) how internet addicted differ from the non-addicted and (2) how these attributes, together with the seductive properties of the Internet, are related to Internet addiction.

Internet addicted are also strongly linked to the pleasure of being able to control the simulated world in online games.

John T. Cacioppo, et al. [14], showed that network linkage data from the population-based Framingham Heart Study can be used to trace the topography of loneliness in people's social networks and the path, through which loneliness spreads through these networks.

Results indicated that loneliness occurs in clusters, extends up to 3 degrees of separation, is disproportionately represented at the periphery of social networks, and spreads through a contagious process. An important implication of this finding is that interventions to reduce loneliness in our society may benefit by aggressively targeting the people in the periphery to help repair their social networks. By helping them, we might create a protective barrier against loneliness that can keep the whole network from unraveling.

To achieve this it is required to make a process of data collection required as a combination of manual and automatic effort in order to produce satisfactory results. Budhaditya Saha, et al. [15], showed that mental illness has a deep impact on individuals, families, and by extension, society as a whole. Co-occurring mental health condition provides the focus for our work on classifying online communities with an interest in depression. Input to the system model is psycho-linguistic features expressed in the posts.

A joint modeling system was formulated in order to classify mental health-related co-occurring online communities from these features. Empirical validation of the model is performed on the crawled dataset where our model outperforms recent state-of-the-art baselines. Result showed the potential of social media and online communities in the early screening and monitoring of mental health-related communities with an interest in depression.

Rohizah Abd Rahman, et al. [16] performed a literature survey of mental health detection techniques for OSN (Online Social Network) users. It was shown that there may be possibility that mental health detection techniques are not limited to one language. However, it was used only Twitter as their OSN platform for data collection instead of other types of OSN.

Helmut Appel, et al. [17], showed that the co-occurrence of depression and envy is both plausible and empirically established. In a quasi-experimental online study, depressed and non-depressed participants indicated their self-esteem and were then presented with specifically set up Facebook profiles that were either attractive or unattractive. The connection between depression and envy was demonstrated with an experimental elicitation of envy for the first time. The results strongly suggest that low self-esteem and consequent feelings of inferiority play a crucial role in depressed individuals' elevated levels of envy. This paper proposes a system for early detection of the condition and referral to the qualified medical staff.

PROPOSED SYSTEM ARCHITECTURE

For early detection of mental disorders we propose a new system called Social Network Mental Disorder Detection (SNMDD) by mining online social data. We formulate the task as classification problem to detect three types of social network mental disorder using machine learning system:

- Cyber-Relationship Addiction, which shows addictive behavior for building online relationships.
- Net Compulsion, which shows compulsive behavior for online social gaming or gambling.

Information Overload, which is related to uncontrollable online surfing. The proposed architecture of SNMDD system for mental health detection consists of several steps such as social network, data extraction, data pre-processing, features selection, data classification using machine learning algorithms and early mental health detection. Figure 17-1 illustrates the proposed architecture of SNMDD system.

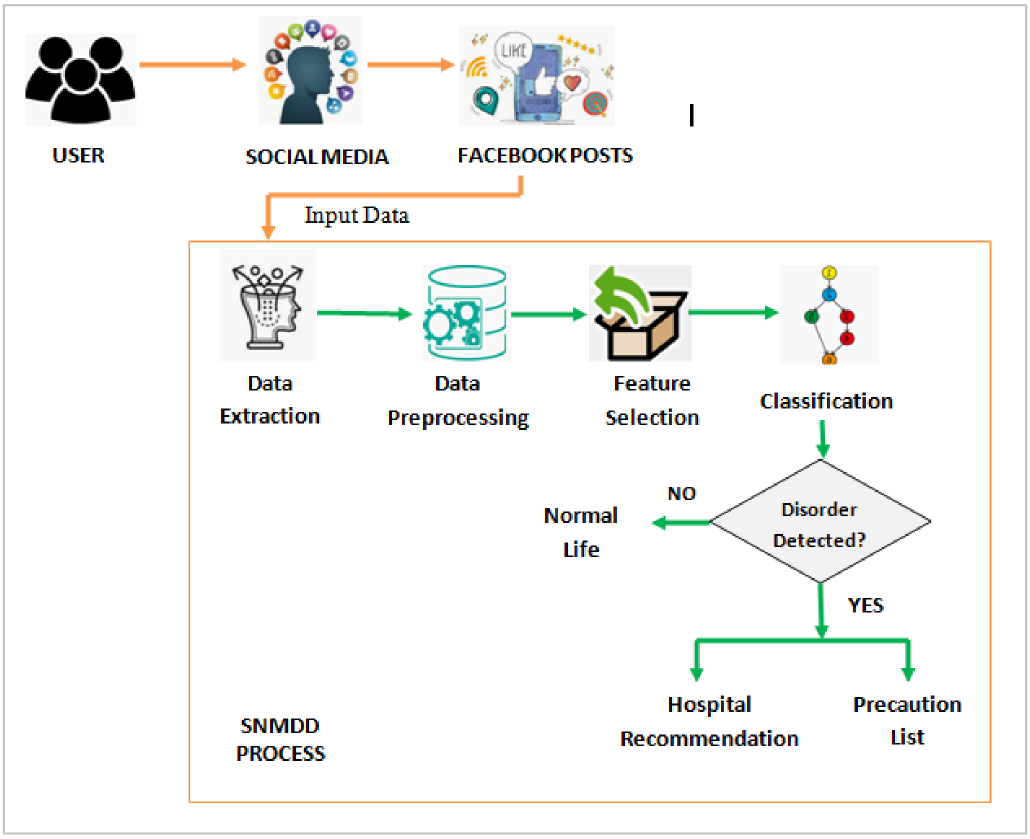

Figure 17-1 System Architecture

In our proposal, we fill the gap of the following constraints:-

- A novel hybrid model - a factor graph model combined with Convolution Neural Network.
- Deep sparse neural network used for trained the data sets
- Machine learning algorithms - ZeroR and Gaussian Processes
- An ensemble of tree classifiers based on a Random Forest algorithm
- Support Vector Machine classifier also implemented for seamless results
- Continuous factor graph model

Steps:

- The users are online social users i.e. Facebook users and the - input to the SNMDD process is the social posts of the users e.g. Facebook posts.
- The SNMDD process is the consists of the users several steps for the detection of mental disorder:-
- Data extraction is the process of data retrieval from various sources. Here SNMDD process extract data from online social media i.e. Facebook data.
- Data preprocessing is the data mining technique which is used to transform raw data into understandable format.
- In this step, the tokenization process is performed to convert sensitive data into unique set of characters that retain all essential information without compromising security.
- The feature selection process is also known as variable selection or attributes selection process.
- This is the process of selecting a subset of relevant features.
- In the SNMDD process the following features are retrieved-
- PR (Personal Relationship), SC (Social Comparison), checkAvgLogin, checkPostLikes, checkFriendRequest, checkTimeSpend, age, prediction, NC (Net Compulsion), CRA(Cyber Relationship Addiction) and IO(Information Overload).
- In ‘classification’ we categorize data into given number of classes. The main goal of a classification problem is to identify the category/class to which a new data will fall under.
- In the SNMDD process, the classification phase categorizes data into two classes i.e. whether user is a normal user or user with mental disorder. A decision tree algorithm is used as classifier. The outcome of the decision is whether mental disorder detected or not detected.
- If mental disorder is detected then user gets recommendation of nearby hospital and precaution list.

REALISTIC METHOD

We formulate the disorder detection task as classification problem.

The general motive of using Decision Tree is to create a training model, which can be used to predict a class or value of target variables by learning decision rules inferred from prior data (training data).

Input: Symptoms Set

Output: Disease Set

Algorithmic Steps:

- ✓ Create dataset
- ✓ Tokenize data
- ✓ Select feature from tokenized data
- ✓ Create root node for the tree
- ✓ If (all inputs are positive, return leaf node positive) If Else (if all inputs are negative, return leaf node negative)
- ✓ Else (Some inputs are positive and some inputs are negative, check condition (Positive¿ negative——Positive ¡negative), then return result)
- ✓ Calculate the entropy of current state H(S)
- ✓ For each attribute, calculate the entropy with respect to the attribute X denoted by H(S, X)
- ✓ Remove the attribute that offers highest value from the set of attributes

- ✓ Repeat until we run out of all attributes or the decision tree has all leaf nodes.

The entropy of the current state H(S)

$$\boldsymbol{H(S) = \sum_{c \in C} - p(c) \log_2 p(c)} \quad \text{---------} \ \textbf{(1)}$$

Where,

S is current data state for which entropy is being calculated.

C is set o classes in S (C = yes or no).

P(c) is the proportion of the number of element in c to the number of elements in set S

Select the attribute which has maximum value of IG(S, X), we use (2):

$$\boldsymbol{IG(A, S) = H(S) - \sum_{t \in T} p(t) H(t)} \quad \text{---------} \ \textbf{(2)}$$

Where,

H(S) is entropy of set S.

T is subset created by splitting set S by attribute A.

$p(t)$ is the proportion of number of elements in t to the number of elements in set S

$H(t)$ Is entropy of subset t.

RESULTS AND DISCUSSION

In the following, we compare the SVM classifier to the supervised learning approach used in SNMDD (i.e. Decision Tree classifier) for single accurate prediction result. Best classifier for any task is it-self task dependent. As shown in Table 17-1, the accuracy of the decision tree classifier is more (i.e. 86%) than support vector machine classifier used in existing system (i.e. 72%).

Table 17-1 Comparison between classification techniques

Sr. No.	Existing System	Proposed System
Algorithm	Support vector machine	Decision Tree
Accuracy	72%	86%

In the proposed SNMDD system as shown in Figure 17-2, the value of precision is 50%, value of recall is 58%, value of F-measure is 53.70% and accuracy is 86% .

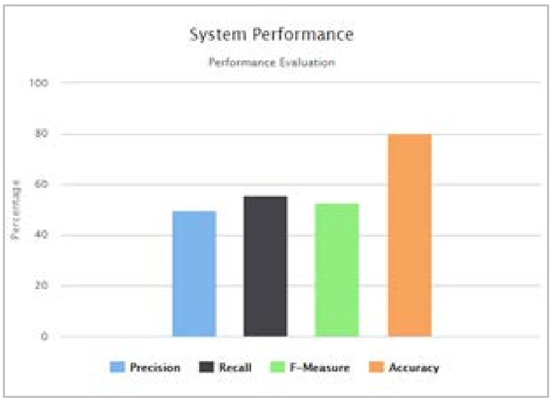

Figure 17-2 System Performance

The system performance was evaluated using the following parameters :

1. Precision =

Precision is the fraction of relevant instances among the retrieved instances

It is the ratio of correctly predicted positive observation to the total predicted positive observations as give by (3):-.

$$\frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \times \mathbf{100} \qquad \textbf{------- (3)}$$

Where, TP is true positive

FP is false positive

2. Recall =

Recall is the fraction of relevant instances that have been retrieved over the total amount of relevant instances. It is the ratio of correctly predicted positive observation to the all observations in actual class –yes.

$$\frac{TP}{TP+FN} * 100 \qquad \textbf{------- (4)}$$

Where, TP is true positive

FN is false negative

3. F-Measure =

F-Measure is a measure of a test’s accuracy. The F-Measure is defined as the weighted harmonic mean of the test’s precision and recall.

F1 score is the weighted average of Precision and Recall.

$$\mathbf{2} * \frac{(\boldsymbol{Precision} * \boldsymbol{Recall})}{\boldsymbol{Precision+Recall}} \qquad \textbf{------- (5)}$$

4. Accuracy =

Accuracy is the most intuitive performance measure and it is simply a ratio of correctly predicted observations to the total observations.

$$\frac{(\boldsymbol{TF+TN})}{(\boldsymbol{TP+FP+FP+FN})} \times \mathbf{100} \qquad \textbf{------- (6)}$$

Where, TP is true positive

FP is false positive

TN is true negative

FN is false negative

In proposed SNMDD system as shown in graph 1, value of precision is 50%, value of recall is 58%, value of F- measure is 53.70% and accuracy is 86%.

The proposed SNMDD system detects three types of mental disorders i.e. cyber relationship addiction (CRA), net compulsion (NC) and information overload (IO). Figure 17-3 shows that among the total users of SNMDD system, 10% of the users are addicted to cyber relationship (CRA), 5% of the users are addicted to a net compulsion (NC) and 11% of the users are addicted to information overload (IO).

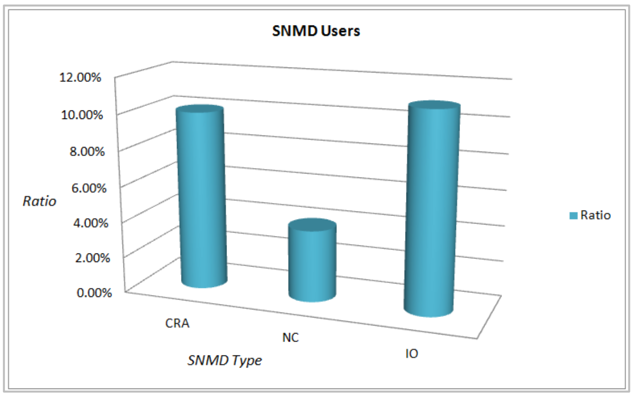

Figure 17-3 Analysis of SNMD Types

CONCLUSION

A system was proposed that consequently recognizes potential online users with SNMDs. Accordingly, we displayed a structure for recognizing clients' Mental Disorder states from clients' per day online networking information, utilizing Facebook post' content just as clients' social associations. Utilizing a genuine internet based life information as the premise, we considered the connection between users' Mental Disorder states and their social collaboration practices.

REFERENCES

[1] Dr.K.Rameshwaraiah, A. Ramakanth. Detecting Stress Based on Social Interactions in Social Networks. @IJRTER-2017

[2] H. Lin, J. Jia, Q. Guo, Y. Xue, J. Huang, L. Cai, and L. Feng. Psy- chological stress detection from cross-media microblog data using deep sparse neural network. In proceedings of IEEE International Conference on Multimedia Expo,2014.

[3] Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. Predicting personality from twitter. In Passat/socialcom2011,Privacy, Security, Risk and Trust, pages 149156, 2011

[4] Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Pentland. Daily stress recognition from mobile phone data, weather conditions and individual traits. In ACM International Conference on Multimedia, pages 477486,2014.

[5] Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, Hiroyuki Ohsaki. Recognizing Depression from Twitter Activity. CHI 2015, Crossings, Seoul, Korea.

[6] Feng-Tso Sun1, Cynthia Kuo1,2, Heng-Tze Cheng1, Senaka Buthpitiya1. Activity-Aware Mental Stress Detection Using Physiological Sensors. MOBICASE 2010, LNICST 76, pp. 211–230, 2012.

[7] Liqiang Nie, Member, IEEE, Yi-Liang Zhao, Mohammad Akbari, Jialie Shen, Member, IEEE,Tat-Seng Chua, Senior Member, IEEE. Bridging the Vocabulary Gap between Health Seekers and Healthcare Knowledge. AUGUST 2013.

[8] K. Young. Internet addiction: the emergence of a new clinical disorder, Cyberpsychol. Behav., 1998.

[9] Yuan Zhang, JieTang, Jimeng Sun, Yiran Chen, and Jinghai Rao. Moodcast: Emotion prediction via dynamic continuous factor graph model. 2013 IEEE 13th International Conference on Data Mining, pages 11931198,2010.

[10] Y. Baek, Y. Bae, and H. Jang. Social and parasocial relationships on social network sites and their differential relationships with users' psychological well-being. Cyberpsychol. Behav. Soc. Netw., 2013

[11] K. Chak and L. Leung. Shyness and locus of control as predictors of internet addiction and internet use. Cyberpsychol. Behav., 2004.

[12] K. Kim, H. Lee, J. P. Hong,M. J. Cho, M. Fava,D. Mischoulon, D. J. Kim, and H. J. Jeon. Poor sleep quality and suicide attempt among adults with internet addiction: a nationwide community sample of Korea. PLOS ONE, 2017

[13] L. Leung. Net-generation attributes and seductive properties of the internet as predictors of online activities and internet addiction. Cyberpsychol. Behav. Soc. Netw., 2004.

[14] J. Cacioppo, J. Fowler, and N. Christakis. Alone in the crowd: themstructure and spread of loneliness in a large social network. J. Pers. Soc. Psychol., 2009.

[15] B. Saha, T. Nguyen, D. Phung, and S. Venkatesh. A system for classifying online mental health-related communities with an interest in depression. IEEE Journal of Biomedical and Health Informatics, 2016

[16] Rohizah Abd Rahman, Khairuddin Omar, Shahrul Azman Mohd Noah and Mohd Shahrul Nizam Mohd Danuri. A Survey on Mental Health Detection in Online Social Network, 2016.

[17] H. Appel, J. Crusius, and Alexander L. Gerla. Social comparison,envy, and depression on facebook: a study looking at the effectsof high comparison standards on depressed individuals. Journal of Social and Clinical Psychology, 2015.