

Research Article

Social Mental Disorders Detection using Machine Learning Approach

Miss. Shilpa M. Kedari and Prof. Pooja Patil

Department of Computer Engineering, D. Y. Patil College of Engineering, Ambi, Talegaon, Pune

Received 10 Nov 2020, Accepted 10 Dec 2020, Available online 01 Feb 2021, **Special Issue-8 (Feb 2021)**

Abstract

The development in social network communication prompts the dangerous utilization. An expanding number of social networks mental disorders such as the dependence on the cybernetic relationship, the over-burden of data and the constriction of the network, have been noticed recently. Currently, the symptoms of these mental disorders are passively observed, which causes late clinical intervention. In this paper, argue that the mining of online social behavior offers the opportunity to actively identify the mental disorders at an early stage. It is difficult to detect mental disorders because the mental state cannot be observed directly from the records of online social activities. Our approach, new and innovative for the practice of mental disorders detection it is not based on the self-disclosure of these mental factors through questionnaires psychology. Instead, we propose a framework of machine learning or the detection of mental disorders in social networks, which exploits the features extracted from social network data to accurately identify potential mental disorders cases. To increase the accuracy of model, we further improve efficiency with performance guarantees. Our framework is evaluated through a user study with no of users of the network. We perform a feature analysis and also apply mental disorders in large-scale data sets and analyze the characteristics of the three types of mental disorder.

Keywords: Online social network (OSN), mental disorder detection, feature extraction, Random forest Classifier

Introduction

With the touchy development in popularity of long range informal communication also, informing applications, online interpersonal organizations (OSNs) have turned into a piece of numerous individuals' day by day lives. Most look into on interpersonal organization mining centers around finding the information behind the information for enhancing individuals' life. While OSNs apparently extend their clients' capacity in expanding social contacts, they may really diminish the up close and personal relational communications in reality. Because of the pandemic size of these marvels, new terms for example, Phone Snubbing and No Cell Phone Phobia have been made to depict those who can't quit utilizing versatile long range informal communication applications.

Social psychological studies have made two interesting Number of people will share their every day events and moods, and interact with friends through the social networks. We can classify using machine learning framework. Due to leverage both Facebook post content attributes and social interactions to enhance mental disorder detection. After getting disorder level, system can recommended user hospital for further

treatment, we can show that hospital on map and system also recommended to take precaution for avoid disorder.

II. LITERATURE SURVEY

This study[1] demonstrates that the linguistic features and topics discussed amongst the online communities have the potential to capture the mental status and presence of mental Health related communities.

Advantages: 1. This framework in order to classify mental health-related co-occurring online communities from these features 2. This framework find 13 types of depressions using social post data.

Disadvantages: Accuracy low

In this paper[2], we proposed a novel data collection mechanism, Subconscious Crowdsourcing, to gather patient and nonpatient Datasets. We then adopted and built our own linguistic and behavioral features to train classifiers for the purpose of mental disorder detection.

Advantages: 1. They propose model that leverage language and behavioral patterns, used particularly in social media, to determine whether a user is suffering from two cases of mental disorder. 2. They implement feature extraction algorithms which aim at removing and assessing the selection bias problem prevalent in mental disorder detection.

Disadvantages: 1. Only used specific language patterns. We have[3] Studies about Daily stress recognition from mobile phone data, weather conditions and individual

traits. Advantages:1. That day by day stress can be dependable received in view of behavioral measurements, got from the client's cell phone action what's more, from extra markers, for example, the climate conditions (information relating to short lived properties of the condition) and the identity attributes. Disadvantages:1. In work environments, where stress has become a serious problem affecting productivity, leading to occupational issues and causing health diseases. They [4] presented a streaming algorithm of NMF that can handle a Twitter stream. Our proposed algorithm can simultaneously Perform real-time topic detection and filtering of outsiders and MG-phrases. Advantages:1. They integrate both the extraction of meaningful topics and the filtering of messages over the Twitter stream. Disadvantages: Accuracy is low because use bag-of-word methods. They studied [5] the interaction between users in the Facebook online social network. They collected data on Friendship links and the activity data based on wall posts for a large subset of the Facebook New Orleans network. Advantages:1. They study the evolution of activity between users in the Facebook social network.2. They studied user interaction between users. This study[6] has inherent biases present in its methodology by utilizing an expedient and convenient self-selected group of Internet users. Advantages:1. They find internet addiction of users 2. Find 5 types of addiction Disadvantages:1. Use limited dataset. The goal of this study [7] was to test the effect of SNS use on users' psychological well-being, in terms of subjective loneliness, interpersonal trust, and SNS addiction. Advantages:1. This study distinguishes types of SNS relationships, and investigates their relationships with social isolation, interpersonal trust, and SNS addiction.2. Find two types of relationships i.e. social relationship and Para social relationship. In this paper[8], to explore social network structures to help sentiment analysis, represents an interesting research direction in social network mining. Advantages:1. Propose to improve sentiment analysis by utilizing the information about user-user relationships. Disadvantages:1. Time consuming process. This paper[9] show depressed users are frequently found other than envy and self-esteem individuals. Advantages:1. This paper find attractiveness, envy, Social Comparison and Depressiveness.2. Use static Data. Disadvantages:1. Limited post observe. This Paper[10] found that frequent social comparisons were associated with a range of destructive emotions and behaviors, including those directed at the self, such as guilt, those directed at others, including lying, and those directed at an outgroup, in the form of in-group bias. Advantages:1. This paper examined the relationship between frequent social comparisons and destructive emotions and behaviors. Disadvantages:1. They cannot address the question of causality. 2. use of self-report to measure social comparison frequency

Proposed Methodology

In proposed system approach, we formulate the task as classification problem to detect three types of social network mental disorder detection using Random Forest Machine learning algorithm:

- i) Cyber-Relationship Addiction, which shows addictive behavior for building online relationships.
- ii) Net Compulsion, which shows compulsive behavior for online social gaming or gambling.
- iii) Over load of data, which is related to uncontrollable surfing.

A. Algorithms

Naive Bayes Classification Algorithm:

1. Given training dataset D which consists of documents belonging to different class say Class A and Class B
2. Calculate the prior probability of class A = number of objects of class A / total number of objects
Calculate the prior probability of class B = number of objects of class B / total number of objects
3. Find NI, the total no of frequency of each class
Na = the total no of frequency of class A
Nb = the total no of frequency of class B
4. Find conditional probability of keyword occurrence given a class:
 $P(\text{value 1/Class A}) = \text{count}/n_i(A)$
 $P(\text{value 1/Class B}) = \text{count}/n_i(B)$
 $P(\text{value 2/Class A}) = \text{count}/n_i(A)$
 $P(\text{value 2/Class B}) = \text{count}/n_i(B)$

 $P(\text{value n/Class B}) = \text{count}/n_i(B)$
5. Avoid zero frequency problems by applying uniform distribution
6. Classify Document C based on the probability $p(C/W)$
 - a. Find $P(A/W) = P(A) * P(\text{value 1/Class A}) * P(\text{value 2/Class A}) * \dots * P(\text{value n/Class A})$
 - b. Find $P(B/W) = P(B) * P(\text{value 1/Class B}) * P(\text{value 2/Class B}) * \dots * P(\text{value n/Class B})$
7. Assign document to class that has higher probability.

B. Architecture

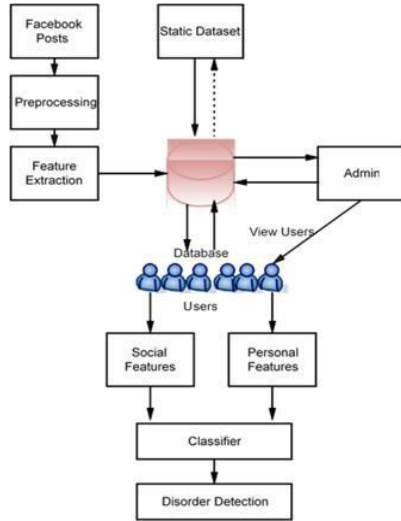


Fig. 1. System Architecture

C. Software Requirements and Specification

The Introduction of software requirements specification provides an overview of all software used in Projects which used the operating system window 7,8,10. The Language used to implementation is java which required the JDK (Java SE Development kit) JDK have many version such as the 1.2, 1.3 and up to 1.8. Platform which used for JDK is eclipse, eclipse have lost of the version. To run the code in eclipse required the Server as the Apache tomcat .Data base used as the MYSQL version 5.4.

D. Mathematical Model

Preprocessing and Feature Extraction:

1. Social Comparison based features: user j, the weighted number of positive newsfeeds based on similarity can be derived as,

$$\frac{\sum_{i \in N(j)} [S(i, j) N_p(i, j)]}{\sum_{i \in N(j)} S(i, j)}$$

where, N(j) is the set of neighbors of user j.

2. Social structure based features

The following minimization problem is formulated to detect structural hole spanners.

$$\begin{aligned} \min & FFD-1AF2,1 \\ \text{s.t.} & FTF = Im, \end{aligned}$$

3. Social diversity based features

Specifically, the diversities of nationality, racial, ethical, religious, and education can be extracted as social diversity based features with Shannon index H as the

diversity index,
 $H = - \sum N_t \text{pilnpi},$

4. Parasocial relationship

The feature of parasocial relationship is represented as,
 a_{out}
 a_{in}

where,
 a_{out} and a_{in} denote the number of actions a user takes to friends and the number of actions friends take to the user.

Results and Discussions

Experimental evaluation is done to compare the proposed system with the existing system for evaluating the performance. The simulation platform used is built using Java framework (version jdk 8) on Windows platform. The system does not require any specific hardware to run; any standard machine is capable of running the application. The experimental result evaluation, we have notation as follows:

TP: True positive (correctly predicted number of instance)

FP: False positive (incorrectly predicted number of instance), TN: True negative (correctly predicted the number of instances as not required)

FN false negative (incorrectly predicted the number of instances as not required),

On the basis of this parameter, we can calculate four measurements

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{F1-Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

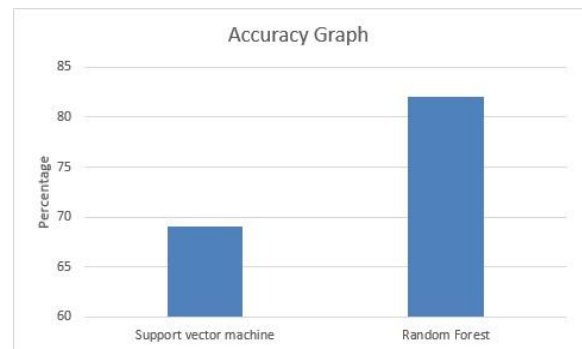


Fig. 2. Graph

Table 1:Comparative Result

Sr. No.	Existing Sys- tem	Proposed System
1	69%	82%

Acknowledgement

Express my true sense of gratitude, sincere and sincere gratitude to my guide to the project Prof. Dr. Pooja Patil for her precious collaboration and guidance that he gave me during my research, to inspire me and provide me with all the laboratory facilities, This it allowed me to carry out this research work in a very simple and practical way. I would also like to express my thanks and thanks to our coordinator ME, Prof. Pooja Patil and Director Dr. Abhay Pawar and all my friends who, knowingly or unknowingly, helped me during my hard work.

Conclusion

In this paper, automatically identify potential online users with SNMDs. Psychological Mental Disorder is threatening people's health. It is non-trivial to detect Mental Disorder timely for proactive care. Therefore we presented a framework for detecting users' psychological Mental Disorder states from users' monthly social media data, leveraging Facebook post ' content as well as users' social interactions.

Employing real- world social media data as the basis, we studied the correlation between users' psychological Mental Disorder states and their social interaction behaviors we recommended the user for health consultant or doctor. We show the hospitals for further treatment on a graph which locate shortest path from current location user to that hospital. We recommended the user for health precaution send on mail for user interaction purpose.

References

- [1] B. Saha, T. Nguyen, D. Phung, and S. Venkatesh. A framework for classifying online mental health-related communities with an interest in depression. *IEEE Journal of Biomedical and Health Informatics*, 2016.
- [2] Chun-Hao Chang, Elvis Saravia, Yi-Shin Chen "Subconscious Crowd- sourcing: A Feasible Data Collection Mechanism for Mental Disorder Detection on Social Media" 2016.
- [3] Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Pentland "Daily stress recognition from mobile phone data, weather conditions and individual traits" pp.2014.
- [4] Kohei Hayashi Takanori Maehara Masashi Toyoda Ken-ichi Kawarabayashi "Real-Time Top-R Topic Detection on Twitter with Topic Hijack Filtering", 2015.