Symptomatic Diagnosis and Prognosis of Psychiatric Disorders through Personal Gadgets

Vidhi Jain
Birla Institute of Technology and Science, Pilani Campus
Pilani, Rajasthan 333031, India
f2014113@pilani.bits-pilani.ac.in

Prakhar Agarwal
Birla Institute of Technology and Science, Pilani Campus
Pilani, Rajasthan 333031, India
f2012277@pilani.bits-pilani.ac.in

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).
CHI'17 Extended Abstracts, May 6-11, 2017, Denver, CO, USA.
ACM ISBN 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3048417

Abstract

Mental disorder has been shrouded as a stigma and disregarded as a secondary issue to physical health. It has become a major contributor to morbidity, disability and at times, fatality. Through our research, we show that the data generated via our daily interactions with technology has consistent patterns to identify symptoms in prodromal phase of degrading mental health.

We propose a methodological data driven system that will help in raising an early alarm on the onset of symptoms of potential psychiatric disorders. The system collects the user's data from different human-computer interfaces to create a fine-grain electronic health portfolio, which can assist doctors in differential diagnosis as well as prognosis.

Author Keywords

Data Collection and Processing; Mental Health Symptoms; Design Technique;

ACM Classification Keywords

J.3 [Life and Medical Sciences]; H.5.m [Information interfaces and presentation]; I.2.6 [Learning]; I.2.7 [Natural Language Processing]

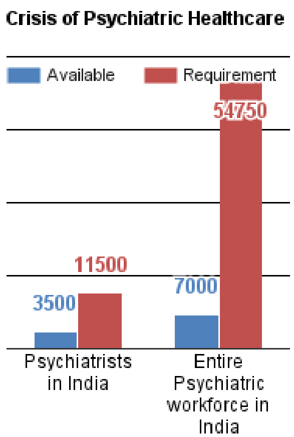

Figure 1: According to Union Ministry of Health and Family welfare in India, there are just 3500 psychiatrists for a total of 1.3 billion people. The entire workforce comprising of clinical psychiatrists, psychologists, psychiatric social workers and psychiatric nurses is nearly 7000 while the actual requirement is around 54750 [10].

Research Questions

How to utilize the rapidly evolving ecosystem of human computer interaction to detect and understand symptoms of psychiatric disorders?
How should we structure a scalable, secure and generalised framework?

Introduction

Mental health is an integral and essential component of an individual's health. We studied the current scenario of mental healthcare. In U.S., one in every five adults (i.e. 43.8 million or 18.5%) experiences mental illness in a given year as per National Institute of Mental health [3]. The situation is significantly worse in developing countries (see Figure 1).

International standard classification tools exist, namely Diagnostic and Statistical Manual of Mental Disorders (DSM-5) by the American Psychiatric Association (APA) and International Statistical Classification of Diseases and Related Health Problems (ICD-10) by the World Health Organization (WHO) for diagnosis of psychiatric disorders. But current practice of diagnoses is often largely subjective rather than being evidence based [11]. To support the practice of evidence based psychiatry, we need exhaustive data about the patients, which may reveal indicative prodromes or patterns of symptoms (see Research Questions).

In our approach, we have first, identified different interfaces where data is generated upon interaction of humans with technology. Second, we developed a scalable Methodological framework for Emotional Journal (MeEJ).

The data from different sources is combined to analyze the history of emotions, behavior and experiences. The model learns patterns in data and analyses whether it is different from that of a neurotically normal person. If it indeed differs, it will be a timely alert about the onset of symptoms of potential psychiatric disorder(s).

Related Work

Research has focused upon automating the process of psychiatric diagnosis since several decades [7]. The process followed by practitioners typically involves an interview, known as a mental status examination, where evaluations are made of appearance, behavior, self-reported symptoms, mental health history, and current life circumstances (see Figure 2).

Some hospitals utilize the assistance of clinical decision support systems [2]. Speech-based psychosis detection [1], emotion and disposition recognition [12] and a variety of computational psychiatry approaches have been studied. Computational tools have been used to identify the first episode of psychotic disorder (FEP) [5]. These techniques remain relatively confined to highly specialized clinics and hospitals, and are not yet inaccessible for majority of the population.

Accuracy of classification becomes critical to ensure reliability of the automated data driven diagnosis compared to a human psychiatric practitioner. Statistical studies such as [9] require sufficient data-set about symptoms to build a robust classification model(s). Web search logs are being explored to identify behavioral patterns which may indicate health conditions of the users [8]. Cognitive healthcare systems like IBM Watson (http://www.ibm.com/watson/health/) are current revolutionary technologies which provide a holistic view of a person's health.

Motivation

Psychiatric problems are often neglected, primarily due to lack of awareness, social stigmas and inaccessibility to mental care services. Some of the major issues are discussed as follows.

Acknowledge the Problem

A person suffering from psychiatric disorder needs to understand that he/she requires help. Delay in realising and communicating about psychiatric symptoms compounds the problem.

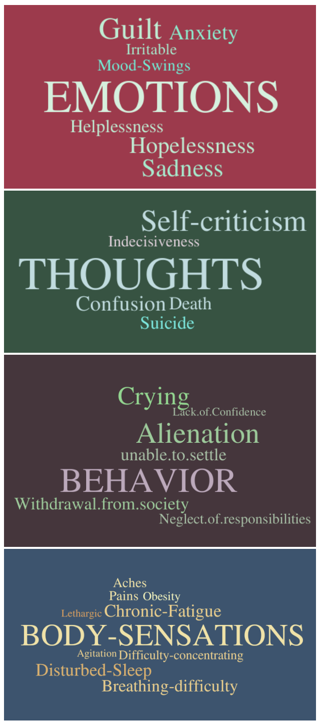

Figure 2: WordClouds depicting how our (a) Emotions (b) Thoughts (c) Behavior and (d) Physical Body Sensations, reveal early symptoms of psychotic disorders

Inaccessibility of Immediate Care
Huge number of untreated cases reflect how inaccessible quality psychiatric services are for a majority cases. Those who are able to seek mental healthcare, often report about “less time spent with doctor" and “heavy medication with side-effects".

Reporting Symptoms and Asking Questions
It is often difficult to share detailed history of all the symptoms for appropriate diagnosis. Social stigmas associated with mental disorders often cause reluctance in visiting a psychiatrist.

Tracking Recovery
It is observed that as soon as the patients start to recover, they tend to give up routine-check visits to their doctor. This may cause significantly higher risk of recurrence of the disorder.

Discovering Symptomatic Patterns
Disorders often overlap and have correlation, such as in psychiatry and diabetes, or in schizophrenia and bipolar disorders. Data driven psychiatric support system can help to identify such correlations and customize medication.

Data Driven Architecture

We propose a system that collects patients’ data to support doctor consultation. The development of this system is a crucial step as it will shape the interactions between patient-system-doctor(s). We focus upon the design for scalable data transmission through this system. We also discuss HCI concerns, its advantages along with its limitations.

Data Sources

Today, both structured and unstructured data has become pivotal in uncovering patterns for better insights on disease/disorder diagnosis. In order to understand day-to-day thoughts, emotions, behaviors and expressions, we have broadly identified three categories of data sources (see Figure 3 Step 1) namely, periodic, aperiodic and stream-based sources (see Categories of Data Sources).

Data Collection

Data is cleaned and processed partially on the local devices like mobile, personal gadgets or local servers, and then collected to be sent to the centralised cloud servers (see Figure 3 Step 2). In order to ensure scalability and security of the model, we propose a subscription based approach for adding and managing data provenances. To channel the huge data, we use a producer-consumer message queue system, Apache Kafka. It has a capacity to handle around 100,000 requests per second. This process is managed by Apache Zookeeper to ensure load balance and fault tolerance.

Storage and Processing

In the previous section, we discussed about edge level processing and passing the data into message queue. This data is fed into processing cluster in private authenticated cloud storage (see in Figure 3 Step 3). We developed the cluster on top of Apache Storm as it provides linear scalability and is self-maintained by Apache Zookeeper.

Learning and Analysis

Once we model the data, it is used to learn and analyze the patterns. We generate two kinds of reports, firstly, raw data analysis and then, information inferences (see Figure 3 Step 4).

Raw Data Analysis
The data recorded from a variety of sources is structured and visualized. The medical experts may suggest hypothe-

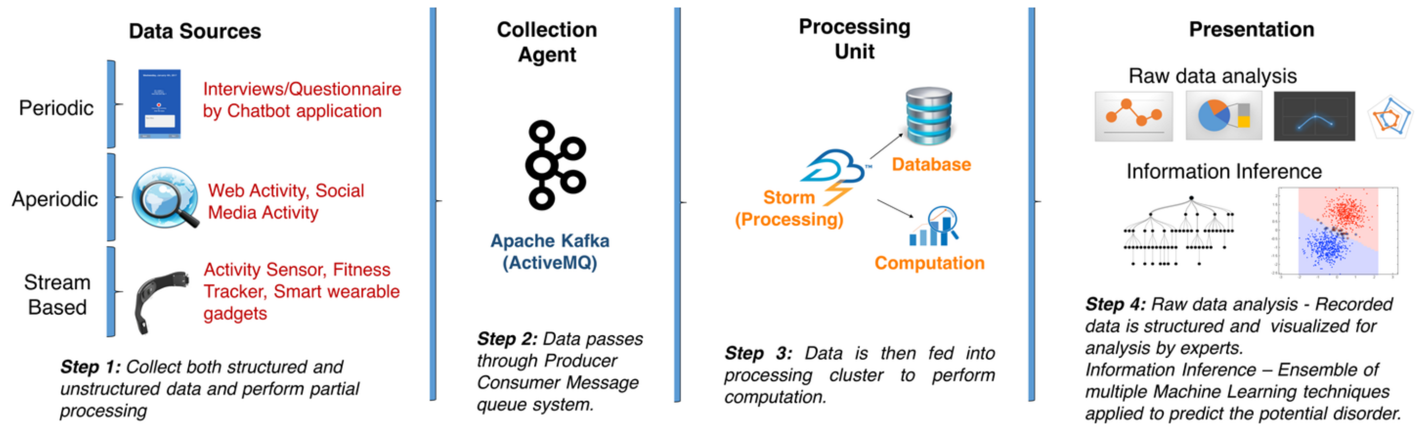

Figure 3: Methodological framework for symptom diagnosis from personal data sources

Categories of Data Sources

Periodic Sources include questionnaires by a chat-bot (see in Figure 4) to regularly record subjective experiences of a person as speech or text. Many existing virtual assistant applications can also be utilized.

Aperiodic Sources are dependent upon the amount of time that a user may want to spend on it. Some of the rich examples are web and social media activity.

Stream based Sources include activity sensors, fitness trackers, and smart wearable gadgets. These are specialized devices to track vital signs of their user in real time, like heart rate, sleep cycle, body movements and composition.

ses by exploring the data that they may not have anticipated in advance and this can lead to interesting follow-up experiments where the hypotheses are formally stated and tested by collecting new data.

Information Inference

The sampled data about the user can indicate several symptoms specific to psychiatric disorders. Data about a particular symptom that has been collected from a variety of sources compensates for some missing values and enhances the confidence level of overall prediction. Several machine and deep learning techniques are applied to the data and then used for ensemble learning to aid prediction of potential psychiatric disorders based upon their symptoms.

Since a basic understanding through emotion and disposition analysis is required, we have worked upon a proof-of-concept based upon Russell's Circumplex model. We utilized the Affective Norms for English Words (ANEW) dataset available for popular words with values for valence and arousal for males and females separately, to train a simple mood classifier as in [6]. For a given a speech or text input, we can visualize the emotional state of the user on a plot (see a testcase for emotion classification and visualization in Figure 5).

Conclusion

We propose to make healthcare 'knowledge rich' by harnessing the data available around us and applying ensemble learning for computational psychiatry to assist diagnosis and prognosis. The key contributions of our work:

- Identification of data sources from day-to-day life to assist evidence-driven symptomatic diagnosis of psychiatric disorders.

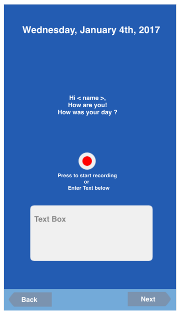

Figure 4: Prototype for Chatbot, a mobile application for recording emotions and experiences via Interviews/Questionnaire.

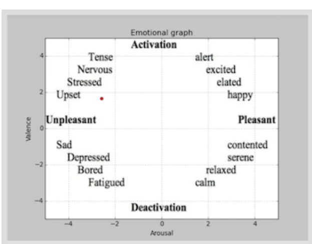

Figure 5: Russell circumplex model. Plot generated for the following test data: *"I am afraid of what others will think of me. I made a horrible mistake. I should never had trusted him."*

- Design of methodological framework for processing of the collected data.

Discussion

The proposed model targets the data sources which the users are already sharing with on-line advertisers and e-commerce businesses. The design of the system focuses on how users can be enabled to record their own health symptoms and at their own discretion.

With more smart devices being developed and connected to the Internet-of-Things, the scope of personal informatics is changing very rapidly. It is observed that people often abandon self-tracking tools [4] and design of such data collection tools needs to consider the psychology of the users very carefully.

Specialized devices would perhaps be designed and prescribed by doctors, especially when they need to track certain symptoms in large number of Clinically High Risk (CHR) patients in their natural routine life [1]. The cases of False Positive inferences need to be carefully examined by expert(s). There is a need for a diagnostic system that comprises of both, tools and psychiatric expertise, to utilize this data effectively.

Advantages

User Control and Freedom Subscription-based model to decide which data sources to allow for diagnostic assessment.

Visibility of Recorded Symptoms Maintain digital health history, visualize data patterns and show it to doctors in case of clinical assessment and, lawyers in case of medical evidence requirement.

Aesthetic and Minimalistic Design Symptoms recorded in current natural setting, through speech input and background processes in most other cases.

Consistency and Feedback Nuanced 24x7 multi-scale time series analysis and subtle feedback through short diagnostic questions.

Disadvantages

Erroneous Recordings Security breaches on social media and erroneous wearable devices can add noise to the recorded data. Sometimes, web activity may not be indicative of the user's state (for example, certain web activity could be for research purposes). Many of the Internet based services have the inherent unreliability in terms of availability and accuracy.

Acceptability This system design is with users' perspective and provides them control over maintaining her/his health logs. It is important on how the psychiatric workforce explores the acceptability of such a system and its results.

Accessibility Contrary to the assumption of regular personal use of smart phones, Internet search, social media or wearable gadgets, many people in developing nations are without these amenities. Some degree of assistance can be provided to them through short targeted diagnostic questions on SMS service.

REFERENCES

1. Gillinder Bedi, Facundo Carrillo, Guillermo Cecchi, Diego Fernández Slezak, Mariano Sigman, Natália B Mota, Sidarta Ribeiro, Daniel C Javitt, Mauro Copelli, and Cheryl M Corcoran. 2015. Automated analysis of free speech predicts psychosis onset in high-risk youths. *Schizophrenia* 1, May (2015), 1–7.

2. Eta S Berner. 2010. *Clinical Decision Support Systems: Theory and Practice* (2nd ed.). Springer Publishing Company, Incorporated.

3. Center for Behavioral Health Statistics and Quality. 2015. Behavioral health trends in the United States : Results from the 2014 National Survey on Drug Use and Health Results from the 2014 National Survey on Drug Use and Health. (2015).
4. Daniel A Epstein, Monica Caraway, Chuck Johnston, An Ping, James Fogarty, and Sean A Munson. 2016. Beyond Abandonment to Next Steps: Understanding and Designing for Life After Personal Informatics Tool Use. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1109–1113.
5. Gorrell et al. 2016. Identifying First Episodes of Psychosis in Psychiatric Patient Records using Machine Learning. *Proceedings of the 15th Workshop on Biomedical Natural Language Processing* (2016), 196–205.
6. C Laurier. 2011. *Automatic Classification of Musical Mood by Content-Based Analysis*. Ph.D. Dissertation. Universitat Pompeu Fabra.
7. Martin Morf, Philip Alexander, and Tom Fuerth. 1981. Fully automated psychiatric diagnosis: Some new possibilities. *Behavior Research Methods & Instrumentation* 13, 4 (jul 1981), 413–416.
8. John Paparrizos, Ryen W White, and Eric Horvitz. 2016. Screening for Pancreatic Adenocarcinoma Using Signals From Web Search Logs: Feasibility Study and Results. *Journal of oncology practice / American Society of Clinical Oncology* 12, 8 (2016), 737–44.
9. Rajamannar Ramasubbu, Matthew R G Brown, Filmeno Cortese, Ismael Gaxiola, Bradley Goodyear, Andrew J. Greenshaw, Serdar M. Dursun, and Russell Greiner. 2016. Accuracy of automated classification of major depressive disorder as a function of symptom severity. *NeuroImage: Clinical* 12 (2016), 320–331.
10. J. Sampathi. 2016. Shortage of psychiatrists hits treatment - Times of India. (2016), 1–2.
11. C. Barr Taylor. 2010. *How to Practice Evidence-Based Psychiatry: Basic Principles and Case Studies*.
12. Jordi Vallverdú and David Casacuberta. 2015. *Ethical and technical aspects of emotions to create empathy in medical machines*. Vol. 74. 341–362 pages.