Master thesis project proposal

Passive Prediction of Mental Health Disorders

Suggested Supervisor at CSE:

Supervisor at Company:

Relevant completed courses:

DIT865, Applied Machine Learning
MSA220, Statistical Learning for Big Data
DIT871, Techniques for Large-scale Data

December 11, 2018

1 Introduction

According to the World Health Organization, major depressive disorder [1] is the largest cause of disability worldwide. Yearly prevalence ranges from 6% to 7% and lifetime prevalence rates are almost double ranging between 15% and 17% of the population [1]. Generalized anxiety disorder [2] is prevalent in a similar proportion of the population, reports suggest yearly prevalence rates around 5% [2]. Both disorder have a strong co-morbidity, and suffer from complications with diagnoses.

One particular issue which contributes to the low diagnostic and treatment rates is how the etiology of both disorders can directly interfere with an individual seeking treatment. Unlike many medical disorders, symptoms of depression and anxiety are not characterized by clear changes in external or physical appearance, as a result detection and diagnosis become more challenging. Individuals are often not aware that their symptoms are due to a psychological disorder and often attribute them to poor mood or external factors [3]. To make this increasingly more complex, depression specifically often also negatively affects an individual's social interactions. This can hamper an individual to actually seek professional support or talk about their experiences. This presents an unique challenge in the psychological community, in how to identify individuals suffering these disorders, even when they may not present themselves for treatment.

2 Problem

One proposed solution to this problem is the concept of passive prediction or diagnose. This solution has come about over the last three or four years, and is not limited to just depression and anxiety. In comparison to the traditional concept of active diagnosis, where an individual suffering certain symptoms would actively seek out a medical diagnosis, this process can now be facilitated by adding a passive element. Unlike a medical professional who has limited time and resources, it is feasible to have machine learning algorithms constantly passively observe an individuals health. Once these algorithms detect certain changes in an individual's health that might be indicative of a disorder, the algorithm can inform the individual and the appropriate human professional for further investigation.

An example of this application is *DeepCare*, developed by a research team at Google [4]. This end to end application is a general approach at diagnosing a wide range of disorders. It has also inspired more specific approaches as well, including that by [5], which predicted suicide attempts up to two years in the

[1] Hereafter referred to as depression

[2] Hereafter referred to as anxiety

future by reviewing psychological records. This field of work allows medical professionals to actively provide interventions to those at high risk before the disorder even sets in.

3 Context

Our contribution to this field can be divided into two areas, a conversational chatbot and a host of machine learning classifiers.

3.1 Conversational chatbot

An increasing number of research studies has demonstrated that in certain situations individuals are more willing to discuss and disclose personal mental health related information during a human computer interaction compared to a human to human interaction [6, 7]. Additionally, the author of [8] have demonstrated that patients react psychologically to conversational agents as if it is human, regardless of whether they think it is human.

Although some work has looked at analysis social media posting for passive prediction, we argue that based on the above research conversational chatbots may have a much stronger chance at individuals expressing behavioural aspects related to depression and anxiety rather then secondary analysis of social media data. In addition, several non profit organizations worldwide have began investigating replacing or improving mental health support call and text lines through the use of dialog agents (Brisbot in Sweden [9], Woebot and Crisis textline in United States [10, 11]).

3.2 Machine learning classifiers

It has been well established within the field of psychology that certain mental illnesses can manifest in the spoken and written language of individuals [12], this concept is termed psycholinguistics [12]. Published work has already shown that machine learning classifiers can be successfully trained to distinguish between people diagnosed with certain disorders and control groups. Some of these disorders are bipolar disorder [13], depression [14] and anorexia [15]. A workshop at the conference and labs evaluation forum in 2017 and 2018 has seen almost 30 publications specially looking at the prediction of depression on social media [15, 16, 17].

4 Goals and Challenges

The overall encompassing goal for this thesis is to explore whether a passive diagnostic conversational chatbot is effective in diagnosing individuals suffering depression and anxiety. We break down the goal into two areas, the conversational chatbot and diagnostic aspect powered by machine learning.

We propose developing a conversational chatbot based on the existing field of research on sequence to sequence neural networks. The chatbot should be able to *appropriately* have a conversation with individuals who may be suffering from depression or anxiety. To the best of our knowledge, this has not been developed before. The main goal will be making the model appropriate for the situation in context. We propose exploring areas like empathy, sensitivity to the participant and preventing foul language. All of which are challenges to the overall goal. The main challenge however will be collection of a suitable dataset, this challenge is explored more in the approach section.

The second aspect of our project is developing a diagnostic aspect to the conversation between the individual and chatbot. Almost exclusively previous work has viewed these disorders as binary classification, either the presence or absence of a disorder at time `x`. However, this offers little value to medical professionals. Rather we propose interrogating our work with the patient health questionnaire (PHQ) psychometric diagnostic scales [18]. These scales identify specific important symptoms and respective cut off for the presence or absents of them. We propose developing separate classifiers to predict a PHQ outcome for both depression and anxiety. By doing so, we ensure our output in an already understandable format for professionals and we can draw on the extensive literature surrounding these scales.

We propose evaluating our work against the questionnaires, to see if text analysis through psycholinguistics and machine learning can offer the same prediction power as self completed questionnaire forms. The challenge in this case will be the level of depth we explore in the development of each classifier. We will want to ensure a model that is best suited to the problem, but is not over-fitted on our dataset.

We would like to acknowledge the following scope however in the goals, we initially propose working exclusively with depression and anxiety. However, we feel that if unexpected challenges occurs we have two possible back up directions. Firstly we could explore another disorder such as anorexia which has both a similar etiology to depression and existing research published on it [19]. Alternatively we could focus more heavily on one aspect of the thesis, either the chatbot or the classifiers.

5 Approach

Again we divide the scope of the project into two areas, the conversational chatbot and the diagnostic aspect.

5.1 Conversation chatbot

The main open source framework for the development of sequence to sequence neural network is OpenNMT [20]. The development primarily requires two things, a dataset in question/answer (QA) format and an appropriate set of hyper parameters for the model. In reference to the challenges mentioned above, there are two methods we can employ to ensure the trained model is appropriate to the situation.

- Option one is to use an well established open source QA dataset about a generic topic and develop a selection of methods that ensure it's appropriateness. Focusing heavily on factor such as empathy, sensitivity to the participant and preventing foul language.

- Option two is to acquire a dataset that is extracted from existing conversations between a depressed individual and a psychologist. We would assume a trained psychologist would already consider factors such as empathy, sensitivity to the participant and preventing foul language. Although it may be more complex to acquire this dataset, we expect to not have to spend as much on developing the aforementioned factors.

The second aspect of this subsection will be choosing a appropriate set of hyper parameters. Given training time for these models is normally within the period of 20 to 30 days, we expect to conduct quite a detailed literature review initially to ensure we do not have to repeatably rerun the model.

There is an established set of evaluation metrics for machine translation models, evaluation of QA systems is both a more recent problem and more complex [21]. Hence we do not note any specific metrics at this stage.

5.2 Machine learning classifiers

As mentioned previously, we propose developing a suite of classifiers to investigate individual PHQ symptoms. Table 1 summaries the four classifiers we propose developing, although we acknowledge possible future room for change depending on aspects such as time and data collection. Data collection may use social media scraping or the Stanford DAIC-WOZ dataset [7].

Table 1: Symptoms to be investigated

Questionnaire	Symptom
PHQ - Depression	Depressed mood every day
PHQ - Depression	Loss of interest in previous activities
GAD - Anxiety	Anxiety presence every day
GAD - Anxiety	Uncontrollable worrying

An early step in the project will be picking a suitable classifier algorithm for each symptom. Classifier will either be supervised or unsupervised depending on the availability of data. Depending on the classifier and data a feature set will be create. The classifier will be designed to try match the outcome of the PHQ, by predicting the presence or absence of a symptom on a four point scale.

5.3 Evaluation

Our evaluation approach will be based against the PHQs. We hope to investigate if a user having a short conversation with our system and running machine learning analysis on their texts can provide a similar or better prediction of depression and anxiety than the PHDs. Aspects such as sample size and ethical approval for dealing with a vulnerable population will have to be explored.

References

[1] Klaus P Ebmeier, Claire Donaghey, and J Douglas Steele. "Recent developments and current controversies in depression". In: *The Lancet* 367.9505 (2006), pp. 153–167.

[2] Dirk Weissenborn, Georg Wiese, and Laura Seiffe. "Making Neural QA as Simple as Possible but not Simpler". In: *CoNLL*. 2017.

[3] RH Belmaker and Galila Agam. "Major depressive disorder". In: *New England Journal of Medicine* 358.1 (2008), pp. 55–68.

[4] Trang Pham et al. "Deepcare: A deep dynamic memory model for predictive medicine". In: *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer. 2016, pp. 30–41.

[5] Colin G. Walsh, Jessica D. Ribeiro, and Joseph C. Franklin. "Predicting Risk of Suicide Attempts Over Time Through Machine Learning". In: *Clinical Psychological Science* 5.3 (2017), pp. 457–469. ISSN: 21677034.

[6] Adam S Miner, Arnold Milstein, and Jefferey T Hancock. "Talking to machines about personal mental health problems". In: *Jama* 318.13 (2017), pp. 1217–1218.

[7] Albert Haque et al. “Measuring Depression Symptom Severity from Spoken Language and 3D Facial Expressions”. In: *arXiv preprint arXiv:1811.08592* (2018).

[8] Byron Reeves and Clifford Ivar Nass. *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge university press, 1996.

[9] Rishabh Chawla and J Anuradha. “COUNSELLOR CHATBOT”. In: *Computer Science* 5 (), pp. 126–136.

[10] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. “Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial”. In: *JMIR mental health* 4.2 (2017).

[11] William P Evans, Laura Davidson, and Lorie Sicafuse. “Someone to listen: Increasing youth help-seeking behavior through a text-based crisis line for youth”. In: *Journal of Community Psychology* 41.4 (2013), pp. 471–487.

[12] James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. “Psychological aspects of natural language use: Our words, our selves”. In: *Annual review of psychology* 54.1 (2003), pp. 547–577.

[13] Yen-Hao Huang, Lin-Hung Wei, and Yi-Shin Chen. “Detection of the Prodromal Phase of Bipolar Disorder from Psychological and Phonological Aspects in Social Media”. In: (2017).

[14] Marcel Trotzek, Sven Koitka, and Christoph M. Friedrich. “Utilizing Neural Networks and Linguistic Metadata for Early Detection of Depression Indications in Text Sequences”. In: (2018).

[15] Faneva Ramiandrisoa et al. “IRIT at e-Risk 2018”. In: (2018).

[16] David E Losada, Fabio Crestani, and Javier Parapar. “Clef 2017 erisk overview: Early risk prediction on the internet: Experimental foundations”. In: *Working Notes of CLEF 2017-Conference and Labs of the Evaluation Forum.* 2017.

[17] Idriss Abdou Malam et al. “IRIT at e-Risk”. In: CEUR Workshop Proceedings. 2017.

[18] Alexandra Martin et al. “Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population”. In: *General hospital psychiatry* 28.1 (2006), pp. 71–77.

[19] Pascale Fung et al. “Towards empathetic human-robot interactions”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9624 LNCS (2018), pp. 173–193. ISSN: 16113349.

[20] Guillaume Klein et al. “OpenNMT: Open-Source Toolkit for Neural Machine Translation”. In: *CoRR* abs/1701.02810 (2017).

[21] Siva Reddy, Danqi Chen, and Christopher D Manning. “Coqa: A conversational question answering challenge”. In: *arXiv preprint arXiv:1808.07042* (2018).