Personalized Treatment of Mental Health Problems using Voice Signal Processing

Suresh Kumar Mukhiya

Høgskulen på Vestlandet
Inndalsveien 28, 5063 Bergen, Norway
past@hvl.no

A proposal submitted for the degree of

Doctor of Philosophy (Ph. D.)
in
COMPUTER SCIENCE

HØGSKULEN PÅ VESTLANDET
BERGEN, NORWAY

September 2018

Signed by:

_____   _____
Yngve Lamo, Professor, PhD   Date
Supervisor, Høgskulen på Vestlandet

_____   _____
Violet Ka I Pun, Associated Professor, PhD   Date
Supervisor, Høgskulen på Vestlandet

_____   _____
Fazle Rabbi   Date
Postdoc, Høgskulen på Vestlandet

_____   _____
Suresh Kumar Mukhiya, M. Sc.   Date
Applicant

1   Introduction

According to WHO (World Health Organization), one in four people in the world will be affected by mental or neurological disorders at some point in their lives. Around 450 million people currently suffer from such conditions [22]. According to Stephens-Reicher, Metcalf, Blanchard, Mangan and Burns [21], it is estimated that about 2.6 million people in EU are suffering from mobility issues and half of these people requires assistive technology so that they can perform an everyday task. Also, about half of the Norwegian population will have experience of mental health problems during the course of their life, and about one-third during the course of one year [20]. These statistics show the incidence of significant growth of mental health disorders but lack of ICT (Information and Communication Technology) system for standardized screening, assessment, and treatment of such mental health problems.

The rising development seen in the use of ICT and the adoption of the Internet facilities among middle and low-income countries have given an opportunity to avail the facilities by paying lesser cost [10]. The use of the Internet has enabled the reach of individuals towards receiving the support and treatment for their mental illness and disorders from multiple means. It made easier for people to seek help through online means mainly by searching about their disease and knowing ways through which they can overcome their mental illness. Reports highlighted that the subscription of mobile phones exceeded up to 80% of the entire population within low developed economies such as Central America, Africa, and South Asia[14]. This phenomenal growth of ICT and lack of ICT systems in mental health assessment and treatment are the driving force of this research project.

The primary objective of the research project is to investigate, evaluate and develop MDSE framework that can be used in assessment of mental health problems using voice signal processing. Section 2 discusses about various motivational aspects of the project. Section 3 summarizes various research questions associated with the project. Afterwards preliminary study section 4, expected outcome section 5, state of art section 6, privacy policy section 7, methodology section 8, progress plan section 9 and conclusions will be introduced.

2   Motivation

The inadequate provision of interventions for the prevention and treatment of mental health problems has become a global challenge across healthcare systems. With a significant increase in mental health problems, it is challenging yet technical and economic pressure to develop an adaptive system to bridge treatment gaps, reduce waiting times for patients and provide interventions at the reduced cost. The other main motivational factors of this research are discussed below:

2.1   Lack of Adaptive Systems

While there has been a handful of researches accomplished on the problem domain, most of the works done are only limited on research papers or in the form of prototype or with very limited accessibility and availability. Most of the Voice Stress Analysis (VSA) system like Psychological Stress Evaluator[1] (PSE) and VSA Mark 100[2] available in the market are used in the form for Polygraph and they are expensive for normal use. The researches done in the domain has not been successfully transformed into Information and Communication Technology systems that can be used in the assessment and treatment of mental health problems.

2.2   Accessible and Available system for self-diagnosis

It is very difficult to provide mental health solutions to each and every people with ease. Moreover, people suffering from mental health requires continuous screening and treatment. It is expensive in one hand and there is a lack of technical resources and medical resources in other. If there is an adaptive system that can be used in the personalized treatment of mental health, it can be accessible and available to everyone.

2.3   Tools for General practitioner

These systems can be used as tools for General Practitioner (GP). In Norway, it is common to

---

[1] http://www.dektorpse.com/products/
[2] https://www.spyzone.com/ccp0-display/liedetectors.html

have a General Practitioner to discuss any kinds of health issue including mental health problems. And it is very difficult to consult an expert in the domain unless one is suffering from a chronic or is in a serious state. In such scenario, adaptive system that can be used in the treatment of mental health problem can be a tool for GP and for people with low, medium mental health issues.

2.4   Self-aware treatment

These adaptive systems can be used for self-aware treatment to monitor one's behavior, stress level and other medical data.

3   Research Questions

The primary aim of this research is to come up with adaptive technology framework that can be used in clinical setup for personalized assessment and treatment of mental health problems. This Ph.D. project focuses on using voice signal to diagnose mental health problems. With this goal in mind, concrete research focus and questions are proposed below. Each question is based on the different stage of the research process, following an incremental development approach. The main hypothesis of this research is: *Voice signal processing can be used in personalized assessment and treatment of mental health problems.*

RQ1 What kind of features, voice signal parameters are taken into considerations in state of art research in the detection of mental health problems? How can these features be categorized?
   1. What voice features can be considered for mental health problem detection?
   2. Can we classify these features based on their performance?

RQ2 What are the approaches and methodologies used in data collection from the mental health patients? How can these methods be applied to clinical diagnosis?
   1. How can we collect data for the empirical purpose? How can we setup infrastructure that can be used in clinical setup?
   2. How can these data be sent to a server? What pre-processing can we perform on the terminal devices like mobile and tablets before sending these data to the server? What are the protocols for storing, encrypting and storing them?

   3. How can we store the data without violating privacy policy? Can we encrypt or compress the data set in this scenario?
   4. How can we access this data in the server? What type of visualization can be done on the server and how processed data be shared with others?

RQ3 What are the state-of-art concepts, methods, algorithms, and technology applied in the diagnosis of mental health problems? How is the performance of these algorithms or ICT technology?
   1. What are the common classification algorithms that can be used in detecting mental health problems?
   2. How can we categorize these algorithms based on their performance or type of mental health problems?

RQ4 Is it possible to create a framework using MDSE that uses voice signal to diagnose mental health problems and provide personalized treatments? Can the framework be utilized in a clinical setup?
   1. Is is possible to create a prototype to show an adaptive system that can be used in the treatment process?
   2. Can the adaptive system be used in clinical setup?

4   Preliminary Study

The author performed a preliminary study to find adaptive technology in the form of mobile applications or web applications that can be used in the treatment of mental health problems. The study, published in the *Journal of Medical Internet Research*, says there are roughly 40,000 health apps available for smart-phone and there are around 800 apps oriented to medical health [7]. According to Anthes, Emily [4] these mobile health apps serve as a digital lifeline and put a portable therapist in every pocket.

   Most of these applications can be roughly categorized into the following groups:

– **Cognitive Remedial Therapy (CRT)**: These type of app targets low depressed people or having a medium level of mental health problems. These people have to go through a series of sessions where in each session they get lessons and a series of tasks. The performance of users is recorded and evaluated over time.

– **Cognitive Behavior Therapy (CBT)**: It takes a hands-on, practical approach to study thinking or behavior that are behind people's difficulties and using the information to change their thought behavior. According to Nilly Mor [18] typical course of CBT consists of three phases. The first phase of treatment focuses on symptom relief. The middle phase addresses cognitive changes and the final phase addresses on maintenance of treatment effect and on relapse prevention. *CBT Thought Diary*[3] is one of the apps that comes under this category. *Depression CBT Self-Help Guid* [4] is another one.

– **Standard Questionnaire**: Many of the mobile apps found online are based on some standard's like PHQ-9, MADRS-S, BDI where a score of mental health problems is calculated based on the responses recorded by the user. *Depression Test* [5] uses PHQ-9 to detect if a person has stress or not. *Moodpath - Depression & Anxiety Test* [6] is another android app that is based on set of questionnaire.

– **Others**: There are other applications found that claim in the treatment of mental health problems using other parameters like communication, awareness creation through online courses, exercises, keeping records or their activities and many more. Table 1 shows the list of applications under consideration for study during preliminary study.

A mind map was created during proposal creation and can be found here [7].

---

[3] https://itunes.apple.com/us/app/cbt-thought-diary/id1010391170?mt=8
[4] https://play.google.com/store/apps/details?id=com.excelatlife.depression
[5] https://itunes.apple.com/us/app/depression-test/id666436210?mt=8
[6] https://play.google.com/store/apps/details?id=de.moodpath.android
[7] https://www.mindomo.com/mindmap/e1b04770ada147519c70432421eb7035

App Name	App Links	Category	Domain
Moodpath - Depression & Anxiety Test	https://goo.gl/gv7DLx	Android, IOS	Tracks psychological, emotional and physical well being records
Depression Test	https://goo.gl/U8Sxy1	Both IOS and Android	Test based on PHQ-9
Depression & Anxiety Self-Test	https://goo.gl/9FCdLd	Android	CBT
Depression Screening Test	https://goo.gl/DtF3Np	IOS	PHQ-9
iCBT	https://goo.gl/exmxPK	IOS	CBT
Social Anxiety Test	https://goo.gl/CYvwsd	IOS	Social Phobia Inventory (SPIN) Questionnaire
Bipolar Disorder Test	https://goo.gl/a2zQFV	IOS	MDQ Screening Questionnaire
Happify: for Stress & Worry	https://goo.gl/h9nFN7	IOS	Activities, Games & Meditation
MoodTools - Depression Aid	https://goo.gl/fV25vz	IOS, Android	Record activities, Videos, lessons, Plans
Depression Test	https://goo.gl/kUoBDq	Android	Questionnaire
Eating Disorder Test	https://goo.gl/MFx9Pd	IOS	26 question screening test
Whats Up? - A Mental Health App	https://goo.gl/FMqyQw	IOS, Android	CBT & Acceptance Commitment Therapy
Talkspace Online Therapy	https://goo.gl/AU5vDn	IOS, Android	Paid chat service
TalkLife	https://goo.gl/bk5Wn2	IOS, Android	Chat service
MY3	https://goo.gl/cAZ7u5	IOS, Android	Connect with People in Need
SuperBetter	https://goo.gl/cWWMeL	IOS, Android	Game for PTSD and social anxiety
Headspace	https://goo.gl/MGrQ8D	IOS, Android	Meditation

Table 1: List of Applications under Study

5 Expected Contributions

Once the thesis is done, it is expected to have the following contributions:

5.1 Contribution to the Project

1. A thesis containing research, outcomes, and evaluation of the work done during the project duration
2. An MDSE infrastructure prototype that can be utilized in clinical setup for data collection, prepossessing of voice data maintaining ethics of privacy.

5.2 Research Papers

RP1 A research paper about *Using voice signal for diagnosis of Depression: Systematic Literature review*

RP2 A research paper about *Methodologies for collection of mental health data in the clinical setup*

RP3 A research paper about *Using voice signal processing for detection of mental health problems*

RP4 A research paper about *Multimodal assistive technologies for mental health diagnosis and monitoring*

6 State of the art works

Speech is a very popular feature for stress analysis and has been used in research in many domain [17,13,12,8,9,2,3]. These researches can be categorized under three headings:

6.1 Voice Analysis

Kim and his team tested the feasibility of using an Interactive Voice Response (IVR) phone system to screen postpartum depression among low-income, English and Spanish speaking mothers [15]. Automatic Speech analysis has been used for the assessment of patients with pre-dementia and Alzheimer disease [16] and the authors claim speech analyses could be an additional objective assessment tool. Tuka and his team used both voice and text for detecting depression [1]. The authors utilized 142 individuals undergoing depression screening and modeled the interactions with voice and text features in a Long-Short Term Memory (LSTM) neural network model to detect depression. Dr. John Hansen from University of

Collarado [6] performed an extensive investigation about the feasibility of voice stress analysis from speech and its applicability to be used in Police departments to detect the stress level of the subjects under investigation. The author claimed the state of art Voice Stress Analysis (VSA) detects the stress with better precision than contemporary Polygraph.

6.2 Voice and Video Analysis

With a long term goal of developing assistive technology that can be used in the treatment of mental health problems, several researchers used both voice signals to detect mental health problems. Comparisons on clinical diagnosis of major depression with automatically measured facial actions and vocal prosody in patients undergoing treatment for depression [5,23,13] has been studied that shows the possibility of detecting depression using both voice and video signals.

6.3 Other Physiological Sensors analysis

In addition to using voice and video signal for detecting mental health problems, there have been several studies where other physiological sensors data like ECG and sleeping patterns are considered for the detection of mental health issues. Tiago Moreira [19] researched on using electrodermal activity (EDA), heart rate (HR), various indices of heart rate variability (HRV), blood pressures (BP), muscle tension and respiration for classification and estimation of subjects with the level of stress.

It can be inferred from the state-of-art works that voice has been successfully used to detect the stress level. However, most of the researches done in the domain use similar data sets. The most popular data set for speech analysis among researchers seems to be SUSAS data set [11]. In addition to that, these researches are more in research papers than in adaptive systems that can be used in a real clinical process for screening, accessing and diagnosing mental health issues. *Depression* is the most common mental health issues to be researched on.

7   Privacy Policy

This Ph.D. research project does not involve the collection of personal information or health records. There is no need to report to the Norwegian Data Protection Official for Research (NSD or other) or Regional Committees for Medical and Health Research Ethics (REK). However, when the adaptive system is used in clinical setup, mental patients records including general information and health data will be saved in the server for diagnosis. The system will require to take user consent and report to NSD regarding the use of mental health data.

8   Methodology

Figure 8 illustrates major phases of the research studies. The initial phase is to perform literature review study to understand the state-of-arts work done in the interest domain.
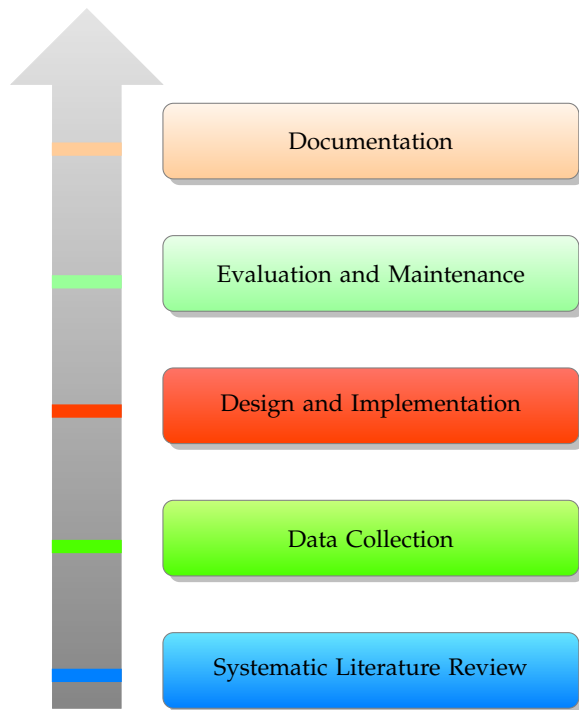
8.1   Systematic Literature Review

A bunch of studies as mentioned in section 6 is done where the voice is considered as parameters for stress detection. However, some authors have considered a few other parameters including video, text and physiological sensors to detect the stress level. In order to get the detailed overview of state-of-art researches on the problem domain, a systematic survey will be done. This will provide the groundwork for answering RQ1 and publishing research paper RP1.

8.2   Data Collection

One of the most important aspects of the research is data collection. Preliminary research showed that there is a lack of training data that can be used training the model. It is one of the important research question RQ2, how data can be collected from mental health patients in a clinical setup. The work will provide grounds for answering research question RQ2 and for publishing research paper RP2.

8.3   Design and Implementation

The author has decided to use Python[8] to design and implement a proof of concept. Any code thus developed will be shared by GitHub[9].

8.4   Evaluation and Maintenance

Once the model is developed, a set of test voice will be used to evaluate the performance of the model. In addition to this, evaluation of data transfer performance, compliance of the system to respect privacy policy, and application of the system in the treatment of mental issues will be done.

8.5   Documentation

The author plans to publish four research papers based on the work done. Some of the targeted conferences where research paper is targeted to

---

[8] https://www.python.org/
[9] https://www.github.org/
[10] https://www.etaps.org/
[11] http://www.modelsconference.org/
[12] https://www.buildstuff.lt/
[13] https://www.journals.elsevier.com/journal-of-logical-and-algebraic-methods-in-programming
[14] http://www.sosym.org/

publish are ETAPS[10], MODELS[11], STUFF[12] and journals like JLAMP[13], SOSYM[14]. A thesis will be produced based on these papers that will work as the documentation of the research done.

9 Progress Plan

Table 2 illustrates overall progress plan of the research for three years.

Year	Semester	Activity
2018	Fall	- Submit research proposal for the admission to the Ph.D. programme. - Take online course on Digital Signal Processing (DSP) on coursera.org - Planning, Analysis and Preparation of Systematic Survey Paper RP1 - Revise basics of Machine Learning and Deep Learning in Python - Participate in design sprint workshop hold in INTROMAT organized by work package two.
2019	Spring	- Course PCS953 Model Driven Software Engineering (10ECTS) - Course PCS902 Research Methodology, Research Ethics and Scientific Work Practice (5ECTS) - Submit Systematic Survey Paper for Review RP1 - Draft and submit research paper RP2 - Attend relevant conferences
2019	Fall	- Take PCS951 Distributed Computing Systems for High Volume Data Processing Course (10ECTS) - Start development of proof of concept - Research and develop MDSE Environment for RQ4 - Continue taking online course - Drafting Research Paper RP3
2020	Spring	- Continue Research and Development - Start writing thesis - Research and develop MDSE Environment for RQ4 - Continue taking online course - Submit research paper RP3 for evaluation
2020	Fall	- Continue with Research and Development - Continue taking online course - Draft and submit research paper RP4
2021	Spring	Write the PhD thesis.

Table 2: Progress Plan for three years

10 Conclusions

This Ph.D. project aims to develop an adaptive system that can be accessible and available for personalized treatment of mental health problems. This research tries to use voice signal from mental health patients and provides methods to detect mental health issues. Based on the fact that voice can be used to detect mental health problems, the research also purposes an adaptive system that can be used in a clinical setup.

The author plans to publish four papers based on the research carried out. These research papers will provide a significant contribution to understanding the problem domain and how an adaptive system can be used in a personalized treatment of mental health issues. In addition to that, the author intends to create an adaptive system that can be used in a clinical setup in Norway. The most interesting contribution of this Ph.D. thesis is its interdisciplinary nature, combining the mental health domain, information and communication technology and opening a broad range for diverse future collaborations.

References

1. AL HANAI, T., GHASSEMI, M., AND GLASS, J. Detecting depression with audio/text sequence mod-

eling of interviews. In *Proc. Interspeech 2018* (2018), pp. 1716–1720.

2. ALGHOWINEM, S., GOECKE, R., WAGNER, M., EPPS, J., GEDEON, T., BREAKSPEAR, M., AND PARKER, G. A comparative study of different classifiers for detecting depression from spontaneous speech. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* (2013).

3. ALPERT, M., POUGET, E. R., AND SILVA, R. R. Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders* (2001).

4. ANTHES, E. Pocket Psychiatry - Mobile Mental-Health Apps Have Exploded Onto the Market, but Few Have Been Thoroughly Tested. *Nature* (2016).

5. COHN, J. F., KRUEZ, T. S., MATTHEWS, I., YANG, Y., NGUYEN, M. H., PADILLA, M. T., ZHOU, F., AND DE LA TORRE, F. Detecting depression from facial actions and vocal prosody. In *Proceedings - 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009* (2009).

6. DARREN HADDAD, SHARON WALTER, R. R. M. S. Investigation and Evaluation of Voice Stress Analysis. *US Department of Justice* (2002).

7. DONKER, T., PETRIE, K., PROUDFOOT, J., CLARKE, J., BIRCH, M. R., AND CHRISTENSEN, H. Smartphones for smarter delivery of mental health programs: A systematic review, 2013.

8. ELLGRING, H., AND SCHERER, K. R. Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior* (1996).

9. FAURHOLT-JEPSEN, M., BUSK, J., FROST, M., VINBERG, M., CHRISTENSEN, E. M., WINTHER, O., BARDRAM, J. E., AND KESSING, L. V. Voice analysis as an objective state marker in bipolar disorder. *Translational psychiatry* (2016).

10. GRAVENHORST, F., MUAREMI, A., BARDRAM, J., GRÜNERBL, A., MAYORA, O., WURZER, G., FROST, M., OSMANI, V., ARNRICH, B., LUKOWICZ, P., AND TRÖSTER, G. Mobile phones as medical devices in mental disorder treatment: an overview. *Personal and Ubiquitous Computing* (2015).

11. HANSEN, J. H., BOU-GHAZALE, S. E., SARIKAYA, R., AND PELLOM, B. Getting Started with SUSAS: A Speech Under Simulated and Actual Stress Database. *Eurospeech* (1997).

12. HANSEN, J. H. L., AND PATIL, S. Speech Under Stress : Analysis, Modeling and Recognition. *Speaker Classification I: Fundamentals, Features, and Methods* (2007).

13. JOSHI, J., GOECKE, R., ALGHOWINEM, S., DHALL, A., WAGNER, M., EPPS, J., PARKER, G., AND BREAKSPEAR, M. Multimodal assistive technologies for depression diagnosis and monitoring. *Journal on Multimodal User Interfaces* (2013).

14. KEMP, S. Digital in 2016, 2016.

15. KIM, H. G., GEPPERT, J., QUAN, T., BRACHA, Y., LUPO, V., AND CUTTS, D. B. Screening for post-partum depression among low-income mothers using an interactive voice response system. *Maternal and Child Health Journal* (2012).

16. KÖNIG, A., SATT, A., SORIN, A., HOORY, R., TOLEDO-RONEN, O., DERREUMAUX, A., MANERA, V., VERHEY, F., AALTEN, P., ROBERT, P. H., AND DAVID, R. Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* (2015).

17. MARTIN HAGMÜLLER, ERHARD RANK, G. K. Can stress be observed by analyzing the human voice? *Signal Processing and Speech Communication Laboratory Graz University of Technology, Austria* (2014).

18. MOR, N., AND HARAN, D. Cognitive-behavioral therapy for depression . *The Israel journal of psychiatry and related sciences* (2009).

19. MOREIRA, T. Estimating Stress Levels with Physiologic Sensors. In *FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO* (2013).

20. SOMMER, M. Mental health among youth in norway. *Norden 69* (06 2016), 7.

21. STEPHENS-REICHER, J., METCALF, A., BLANCHARD, M., MANGAN, C., AND BURNS, J. Reaching the hard-to-reach: How information communication technologies can reach young people at greater risk of mental health difficulties. *Australasian Psychiatry* (2011).

22. WHO. World health report, 2018.

23. YANG, Y., FAIRBAIRN, C., AND COHN, J. F. Detecting depression severity from vocal prosody. *IEEE Transactions on Affective Computing* (2013).