ulm university universität
uulm

Ulm University | 89069 Ulm | Germany

Faculty of Engineering,
Computer Science and Psychology
Institute of Databases and Information Systems (DBIS)

Title of the thesis

Master Thesis in Master Software Engineering at Ulm University

Presented by:
Christoph Meyer
christoph-1.meyer@uni-ulm.de

Examiner:
Prof. Dr. Manfred Reichert
Dr. Felix Beierle

Advisor:
Dr. Felix Beierle

2022/23

Last updated November 7, 2022

Typesetting: PDF-LaTeX 2_ε

Abstract

Contents

1 Introduction

The world health organization (WHO) showed that 20 percent of children and adolescents suffer from mental health conditions. Suicide is the second most death reason among 19-29 years olds[1]. With shortages in the number of therapists, there is a need for computer-assisted psychotherapy (CAT). Self reports and retrospective of habits and feelings are a fundamental concept in improving mental health. For therapist it's the only chance to get an closer insight in the patients behavior during real live scenarios which helps them to find good strategies and treatments. The problem with descriptions of events, feelings or behavior that lie in the past is that they usually do not correspond to reality and vary greatly. Memories are changed by external circumstances. Thus, negative experiences are perceived more strongly than positive ones[**?**]. Ecological Momentary Assessment (EMA) or diary studies address this issue[**?**]. In this type of reports, the patient describe their habits on a daily basis, which results in a less biased retrospective because the events are still present. A additional benefit, Daily journal writing helps patients, to train their own mindfulness and align their own focus on the progress they make. For example Hirage et al. showed that writing a diary can help people after a surgery to set and achieve own goals in their treatment[**?**]. Even for healthy people, journaling can help reduce anxiety and stress and reduce the risk of mental illness. Designing a text offers a cognitive difficulty, making it more difficult to access and integrate into daily routines. Mood Tracker apps address this issue by providing a more easy way to track their emotions on their smartphone. Online Therapeutic tools like "Moodscope" showed scientific proven improvements of the users mental health[**?**]. Therefore, mood trackers are serious category in CAT and 14.2 % of all mobile health applications are mood tracker. Although the increasing number of mood trackers is a good thing, the quality of these apps varies greatly. Scientific reviews of these criticize that many apps are developed without the instruction of psychological professionals and are more in line with the opinions and wishes of users[**?**][**?**]. But what is the opinion of the users? What exactly do users expect from this type of application and how is it implemented? While earlier attempt use randomly selected user reviews[**?**], user interviews[**?**] or the mobile Mobile Application Rating Scale (MARS)[**?**] we want to use an natural language approach to cover those questions. We want to know which issues

[1] `https://www.who.int/news-room/fact-sheets/detail/suicide` Accessed on 2021-11-01

does user have with those applications and if those issues are in common with earlier researches. Also we want to cover, if academic designs for record mood behavior are in common with user practice. Is it more important to track the emotions precisely or is an easy and more accessible representation such as Emojis good enough for user satisfactions?

2 Goal of this thesis

2.1 Problem

Allthough such applications are reviewed by experts we explore some limitations according their methods. Many of the reviews are targeting mental illnesses such as bipolar disorders, depression or anciety as a motivation to use those applications. Despite the fact that this can be an important motivation for some people, it is not the only one. Schueller et al. [**?**] showed in their interview study that many healthy people are using those apps for self awereness after a bad events in their life. They used those apps as a way to reflect on their life and to improve their self esteem. This indicates that the target group of those applications is not only people with mental disorders but also without. Reviews which are targeting mental disorders can miss important aspects of those applications which are important for healthy people. Another limitation is that non of the reviews covering all user reviews. Caldeira [**?**] and Balaskas et al. [**?**] are selecting random reviews and read them by theirself. Although this captures the context of the reviews very clearly, it means that a large proportion of user reviews are ignored. Added interview studies, such as those used by Balaskas [**?**] and Schueller [**?**], improve internal validity but can also be criticized for their external validity. Modern approaches which includes data mining can cover much more reviews but suffer in terms of internal validity. The goal of this thesis is to find a method which covers all user reviews and is able to capture the context of the reviews.

2.2 Research objective

In this study we want to validate the findings of earlier studies based on natural language processing (NLP) appraoches. We want to explore if modern NLP approaches are usable for this analysis and if the findins of earlier studies can be validate or disprove based on our results on all reviews.

2.2.1 Research questions

1. **RQ1:** Can modern NLP approaches be used to analyze user reviews in context of feature importance?

2. **RQ2:** Are the findings of those approaches can be used for validate or disprove the findings of earlier studies?

2.3 Expected results

We expect that with state of the art NLP approaches we are getting a better external validity without loosing .

3 Methology

3.1 Data collection

This study is divided in two separated parts. The first part is a systematic review of the applications. We follow the guidelines of the reporting checklist of the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA[**?**]). We define mood trackers as: "Applications for measuring and reporting mood by themselves on a daily basis". We exclude applications which are:

- main focus is not mood tracking (eg. *Daily Diary: Journal with Lock* [1])
- doesn't collect mood behavior (eg. journalistic [2])
- tracking others behavior such as parenting applications or relationship applications (eg. *behavior diary* [3])
- only available as a web app and not included in any applications store (eg. *moodtracker.com*[4])
- targeting a specific group of people (eg. *Bipolar Mood Tracker*[5])

As data source we are using the google play store[6] as well as the apple appstore because the two operation systems covers more than 99% of the worldwide mobile operation system market share. Our search queries are: ["mood tracker", "mood journal", "mood ema", "emotion tracker"]. For feature extraction we are using the app descriptions as raw data. We include only applications which are available at the time period of our study (November 2022). The User Reviews are crawled with appbot[7].

[1] `https://play.google.com/store/apps/details?id=com.daily.journal.diary.lock.mood.tracker.free&hl=en&gl=US` Accessed on 2021-11-01

[2] `https://play.google.com/store/apps/details?id=com.journalisticapp.twa&hl=en&gl=US` Accessed on 2021-11-01

[3] `https://play.google.com/store/apps/details?id=in.co.skycap.behaviourtracker&hl=en&gl=US` Accessed on 2021-11-01

[4] `https://www.moodtracker.com/` Accessed on 2021-11-01

[5] `https://play.google.com/store/apps/details?id=com.bipolar_flutter&hl=en&gl=US` Accessed on 2021-11-01

[6] `https://play.google.com/store/games?hl=en&gl=US` Accessed on 2021-11-01

[7] `https://appbot.co/` Accessed: 07.11.2022

3.2 Analyzing User Reviews

For analyzing the user reviews we are using two types of data mining. In our first study we are using n-gram frequency to get a first impression on the most important features. In our second study we are using topic modeling to get a better understanding of the context of the reviews.

3.2.1 N-gram frequency

3.2.2 Topic modeling

3.3 Evaluate the results

4 Timeline

List of Figures

List of Tables

Declaration

I hereby declare that this thesis titled:

Title of the thesis

is the product of my own independent work and that I have used no sources or materials other than those specified. The passages taken from other works, either verbatim or paraphrased in the spirit of the original quote, are identified in each individual case by indicating the source. I further declare that all my academic work was written in line with the principles of proper academic research according to the official “Satzung der Universität Ulm zur Sicherung guter wissenschaftlicher Praxis” (University Statute for the Safeguarding of Proper Academic Practice).

Ulm, ...

Christoph Meyer