Thesis Portfolio

An Automated Machine Learning Pipeline
for the Monitoring and Forecasting of Mobile Health Data
(Technical Report)

Confronting the Duality of Using Potentially Harmful Technology
as Host to Beneficial Technology
(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Joseph Keogh
Spring, 2021

Department of Systems and Information Engineering

Table of Contents

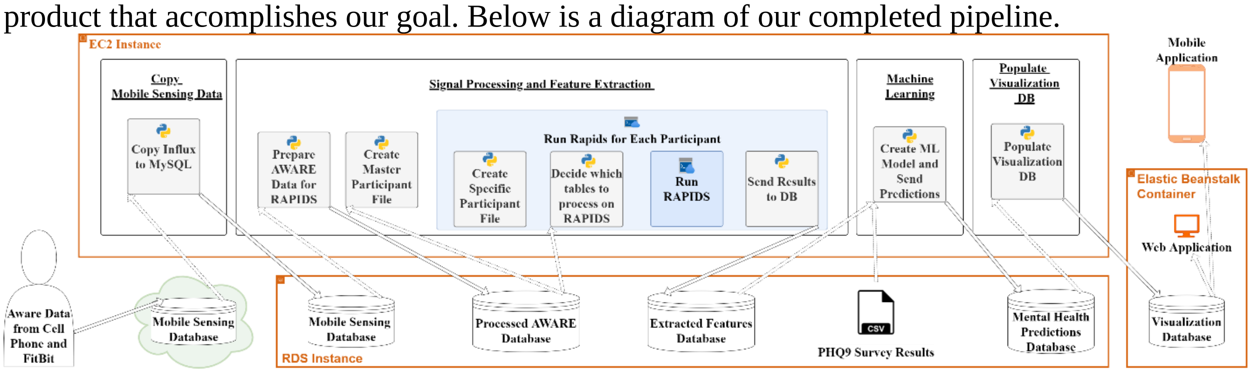
Sociotechnical Synthesis

Introduction

The relation between the STS Research Project and the Technical Project is that the STS Research Project was designed to directly research the ethics involved in the Technical Project. The Technical Project is a mental health application for monitoring and forecasting depression in adolescents utilizing passive sensing and machine learning techniques. The ethical concern is that the application leans heavily on the adolescent's use of their smartphone while knowing cell phones have been shown to increase depression. Creating an ethical duality of putting potentially beneficial technology in a potentially harmful medium.

Technical Project Summary

The Technical Project was to develop an automated pipeline to monitor and forecast mental health data. We specifically implemented our pipeline for the monitoring and forecasting of adolescent depression. Our project consisted of four main sections: data collection, signal processing, machine learning and data visualization. We were successfully able to produce a product that accomplishes our goal. Below is a diagram of our completed pipeline.



STS Research Project Summary

The specific problem that will be addressed in the STS Research Project is whether it is ethical to provide a potentially beneficial technology through a potentially harmful medium? Our specific problem being is it ethical to use a cell phone, which we know can deteriorate the mental health of adolescents, to possibly improve a user's mental health?

This is done mainly by comparing the pros and cons of an individual taking mental health prescription drugs and our application. These two are analogous as they both have the potential to provide mental health benefits with possible negative side effects. We provide an analysis framework for individuals to use to decide whether the Moodring application is right for them. Finally, the research project provides ethical recommendations to ensure the user has the best information to decide on how to best handle their mental health.

Conclusion

Thankfully, the STS Research Project and Technical Project were able to benefit from a symbiotic relationship. The other could not exist in its fullest form without the other. The STS Research Project provides the ethical groundwork for the Technical Project to be feasible in the real world while accounting for the human aspect of users. And the Technical Project provides the main topic of concern to the STS Research.

An Automated Machine Learning Pipeline for the Monitoring and Forecasting of Mobile Health Data

A Technical Report submitted to the Department of Systems and Information Engineering

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Joseph Keogh
Spring, 2021

Technical Project Team Members
Ana Bonaquist
Meredith Grehan
Owen Haines
Neil Singh

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature Joseph Keogh Date May 11, 2021
Joseph Keogh

Approved ______________________________ Date __________
Afsaneh Doryab, Department of Systems and Information Engineering

An Automated Machine Learning Pipeline for Monitoring and Forecasting Mobile Health Data

Anna Bonaquist*, Meredith Grehan*, Owen Haines*, Joseph Keogh*, Tahsin Mullick*, Neil Singh*,
Sam Shaaban†, Ana Radovic‡, and Afsaneh Doryab*
* University of Virginia, Charlottesville, VA
‡ University of Pittsburgh, Pittsburgh, PA
† NuRelm, Pittsburgh, PA
* {acb2qv, mhg8gq, ojh2et, jgk7uf, tum7q, ns9wr, ad4ks}@virginia.edu
‡ ana.radovic@chp.edu
† sam@nurelm.com

***Abstract*—Mobile sensing and analysis of data streams collected from personal devices such as smartphones and fitness trackers have become useful tools to help health professionals monitor and treat patients outside of clinics. Research in mobile health has largely focused on feasibility studies to detect or predict a health status. Despite the development of tools for collection and processing of mobile data streams, such approaches remain ad hoc and offline. This paper presents an automated machine learning pipeline for continuous collection, processing, and analysis of mobile health data. We test this pipeline in an application for monitoring and predicting adolescents' mental health. The paper presents system engineering considerations based on an exploratory machine learning analysis followed by the pipeline implementation.**

I. Introduction

Technology advances and widespread use of smartphones and fitness trackers have made these devices targets for mobile health applications. Passive sensing capabilities embedded in smart devices provide the opportunity to track and monitor behavioral cues related to health and wellbeing.

Mobile health approaches using passive sensing have focused on two main processes, namely representation or modeling of data. Representation refers to the reporting or recording of mental health data in applications, many of which are web or mobile based. Users typically interact with such applications to either record health states or receive treatment for illness (e.g., [1], [2]). Modeling is the act of utilizing data or features to produce insights on a user's mental health via passive sensing. Specifically, modeling approaches focus on the application of machine learning algorithms and their potential role in predicting health outcomes such as depression or anxiety (e.g., [3]–[5]).

However, there is a dearth of work that integrates both the modeling and representation paradigms in an end-to-end product through processing and modeling raw sensor data and communicating the insights to the stakeholders continuously. The focus of this work is on creating such a data pipeline, an automated suite, that will allow for raw mobile device data to predict health outcomes such as depression.

In the following section we describe the essential components of the pipeline and discuss existing tools and approaches to build each component. We then present an exploratory machine learning analysis that informs the implementation. We demonstrate the practicality of the pipeline with a use case on data from a group of adolescents facing depression.

II. Automated Processing and Prediction Pipeline

This section outlines the engineering design considerations for an automated processing and prediction pipeline for mobile health data. In the first part, we describe the architectural components of such a system including:

1) strategies for collection of subjective and objective behavioral data
2) methods for processing and preparing data for modeling
3) machine learning and analysis methods suited for modeling passive time series data
4) strategies for communicating results and insights back to the stakeholders (e.g., patients or clinicians)

In the second part, we demonstrate the feasibility of the automated pipeline through a case study using data from adolescents with depression.

A. Data Collection

1) Passive Sensing: Different mobile data collection frameworks have been developed over the past few years including AWARE Framework [6], Intellicare [7], MindLamp [8], RADAR [9], Sensus [10], CARP [11] and Beiwe [12]. These tools can use both active and passive sensing to collect social, behavioral, and cognitive data from users. The sensor data and derived behaviors from the user's phone and wearable trackers are utilized via a mobile application. Example data sources including GPS, accelerometer, call logs, screen status, wifi, and bluetooth allow inferences on movement patterns, physical activities, social communications, and phone usage. The data is then stored in a database or cloud storage where further analysis can be performed. In this work, we use the AWARE

This project is funded under NIMH project 1R44MH122067-01. *(Anna Bonaquist, Meredith Grehan, Owen Haines, Joseph Keogh, Tahsin Mullick, and Neil Singh contributed equally to the work.)*

Framework, an open-source platform with both Android and iOS clients that has been used in numerous studies to investigate health conditions e.g., [13]–[17]. AWARE Framework also integrates a study module to allow for monitoring of data streams which reduces the implementation efforts of researchers and mobile health developers.

2) Self-reports: Mobile health applications often employ regular subjective assessments from users in form of diaries or surveys. The frequency of self-reports depend on the application and can range from several times a day to once every few months. Many applications use Ecological Momentary Assessment (EMA) [18] to trigger short surveys to collect information from users in certain situations. There are also a number of different software options that allow for online self reporting including survey platforms such as Qualtrics and REDCap. Ideally, mobile applications should be able to objectively assess the state of individuals from sensor data alone without relying on the subjective assessments that are prone to diverge from reality. However, for the purpose of feasibility demonstration, our pipeline allows flexibility for collection of self-report data from different sources. In our implementation, we use data from REDCap, a platform implemented in alignment with most human-subjects research privacy standards, such as HIPAA.

B. Data Processing

Raw data from mobile devices is often noisy and incomplete and needs to be cleaned. Additionally, the raw data is rarely in a usable format for direct interpretation. As such, the pipeline should include mechanisms for cleaning and aggregating the data. Mobile feature engineering [19], [20] has been useful in drawing important insights from data. A few tools, including Digital Biomarker Discovery Pipeline (DBDP) [21], Health Outcomes through Positive Engagement and Self-Empowerment (HOPES) [22], and Reproducible Analysis Pipeline for Data Streams (RAPIDS) [19] have been developed to process mobile data streams. We implement RAPIDS in our pipeline because of its distinct advantages compared to other alternatives. RAPIDS offers an open source modular feature extraction platform originally tailored for data collected with AWARE and Fitness trackers. It also allows for flexible segmentation of time series and supports extraction of over 300 combined features related to movement, physical activities, social communications, phone usage, and sleep, most of which are listed in [20].

C. Modeling

1) Machine Learning: Machine learning models can be applied to behavioral features to predict a health outcome, e.g., [3], [23], [24]. Most machine learning methods are supervised and use patients' self-reported status or the clinical assessment to train algorithms and build models for future prediction of health status. A wide variety of machine learning algorithms can be applied to mobile data for health prediction. In their review of smartphone-based passive sensing applications, Cornet and Holden [23] stated that support vector machines, Bayes classifiers, decision trees, random forests, and linear regression were the most popular machine learning models used for predicting the status of a patient. Meta algorithms such as random forests and gradient boosting combine bagging and boosting to improve the machine learning results. Ghandeharioun et al. [3] demonstrated that such ensemble based methods generalized better on the test set than non-ensemble based methods. For this reason, we choose to try a variety of these algorithms including XGBoost, XGBoostRF, Random Forests, ExtraTrees, Gradient Boosting, Adaboost, Light GBM, and Catboost to determine the best fit.

In addition to learning algorithms, the pipeline should support continuous model building and refinement as new data is added to the streams. Current approaches in this domain aim to automatically produce test set predictions for a new dataset, e.g., AutoML [25], Auto-Weka [26], and Auto sklearn [27]. The training set and testing strategy, however, can have a major impact on the outcome and affect the generalizability of a machine learning model. Doryab et al. [24] discussed two kinds of validation approaches: individual patient models and unified patient models. Individual patient models use only the patient's data to predict their health state, while unified patient models learn from the entire sample of patients to make health predictions. In our implementation, we first run an exploratory analysis of the target dataset (as described in section III-C) to evaluate the training and testing methods and the corresponding machine learning outcomes. This analysis informs the optimal strategies to be implemented in the pipeline.

D. Communication and Visualization

Previous efforts in communicating health data have revealed the importance of evaluating the needs of multiple stakeholder groups and designing to fit those needs. For example, Abdullah et al. [28] uses patients' smartphones to both actively and passively track daily rhythms and to provide effective feedback that can help patients maintain a regular daily rhythm. It also feeds this clinically valuable information back to patients' physicians so that the physician knows how the patient is doing. In the development of Monarca, Frost et al [29] handled the disparate needs of stakeholder groups through the development of separate platforms for communication. A mobile application was developed for patient use, which allowed patients to view their mood level as predicted by the machine learning algorithm, as well as the factors contributing to their mood. A mobile application provides a convenient way for users to view the information that they value quickly, demonstrated in the MoodRhythm application [28]. In the development of Monarca, a web portal was created for use by care providers and researchers which displayed data of multiple patients. This web portal also allowed for care providers to see which patients were most at risk and possibly in need of immediate intervention. For our pipeline, we also designed a mobile application and web portal to inform clinicians and researchers about behavioral features selected by the machine learning algorithm as well as the actual and predicted health outcome (see figure 2). We evaluated the wireframe designs for the mobile application through usability studies with the actual patients. The usability discussion is out of the scope of this paper and is reported elsewhere.

III. Pipeline Implementation for Real-time Assessment of Mental Health in Adolescents

This section details the pipeline components as described in the previous section. We present the implemented architecture (Fig. 1) and modeling informed by the exploratory analysis of a sample dataset from adolescents with depression. The specific components of the implemented pipeline are as follows:

1) **Data Collection**:
    - Passive raw sensor data collected on a daily basis via AWARE and Fitbit are stored in an InfluxDB database.
    - Self-reports PHQ-9 survey for assessing depression are collected once a week through REDCap.
2) **Data Processing**: AWARE data is cleaned and fed into the RAPIDS framework along with a configuration file that specifies the processing settings for feature extraction.
3) **Machine Learning**: Models of depression states are built from extracted features using PHQ-9 scores as ground truth for comparison.
4) **Web Portal and Mobile Client**: The machine learning results including predicted depression score and important features included in the model are communicated on the web portal and mobile clients.

The pipeline is designed to run once a week to align with the frequency of the REDCap surveys. We use Python libraries including pandas, scikit-learn, PyYAML, and SQLAlchemy for development and host the pipeline on an Amazon EC2 instance to promote automation and scalability.

A. Data Collection

AWARE data was originally stored in an InfluxDB database. InfluxDB is based on InfluxQL, a database dialect designed for time series data. Since new data is added to InfluxDB on an ongoing basis, performing further processing task on this database would slow down both data storage and query. As such, we created a secondary database for processing and added the data to this database on a weekly basis. We used a MySQL database in this implementation because of the RAPIDS incompatibility with the InfluxDB format. The duplication of data was done in three steps: copying the schema, copying the data, and verifying the data schema.

In the first step, the schema was copied from the InfluxDB database to the MySQL database. This was done by querying both the InfluxDB 'fields' and 'tags', similar to traditional SQL 'columns' and 'primary keys' respectfully, for all tables in the InfluxDB. The fields and tags were then used to create the respective columns in the MySQL database for each table. This was done to ensure the schema stayed in-tact during the data copying process.

In the second step, the data itself was inserted on a weekly rolling basis. To prevent the overload of the InfluxDB server, queries were sent on a regular basis via Python's 'influxdb' package. Queries were sent by table, by user, and by single day.

In the final step, the schema of the query result was verified. An artifact of InfluxDB is that if all the data for a column is null, the column will be excluded from the query result. To correct this, using the copied schema, we populated any missing columns in our query result with null values. Each query result was then written to the MySQL database via pandas SQLAlchemy and pandas libraries.

For the mental health monitoring of adolescents, the PHQ-9 (Patient Health Questionnaire) [30] survey was implemented on REDCap. PHQ-9 includes nine questions and the scores range from 1 (no depression) to 27 (severe depression). To assess the level of depression, the scores are categorized as minimal depression (1-4), mild depression (5-9), moderate depression (10-14), moderately severe depression (15-19), and severe depression (20-27). The collection of this data was facilitated by the UPMC Children's Hospital of Pittsburgh. As described later, we used both PHQ-9 scores and depression categories in our exploratory machine learning to assess the feasibility of classification vs. regression for our pipeline implementation.

B. Data Processing

The data processing is outlined in the following steps:

1) Data preparation: To prepare the raw mobile data for feature extraction, we first removed redundant columns to match table schemes appropriate for RAPIDS. The modified tables were stored in a separate MySQL database.

2) Create master participant file: We used information from our prepared AWARE MySQL data to generate a master participant file that consolidates information on all participants in the study. This file serves as a reference for RAPIDS. It contains date, participant unique identifier, and device (cell phone) unique identifier information [1].

3) Run RAPIDS to extract features from individual users: RAPIDS is designed to process an entire dataset in one run. Our implementation required ongoing processing and feature extraction for each individual user. As such, we adjusted the settings of RAPIDS to process data from individual users. Our script modifies a centralized configuration file based on data present for each user. To verify a given feature can be extracted, the script queries relevant data tables. The RAPIDS script is then executed to extract weekly features from the prepared AWARE data, and the extracted features for each user are sent to a database for storage and compiling.

4) Obtain PHQ-9 scores: The PHQ-9 scores were calculated and stored in a file on a regular basis. Ideally, the self-report measures are read through the API. However, an issue in matching the timestamp of the survey data prevented us from

[1] We used RAPIDS version 0.4.3. Further information on the current version of RAPIDS, its configuration, and the behavioral features available can be found at https://www.rapids.science/latest/

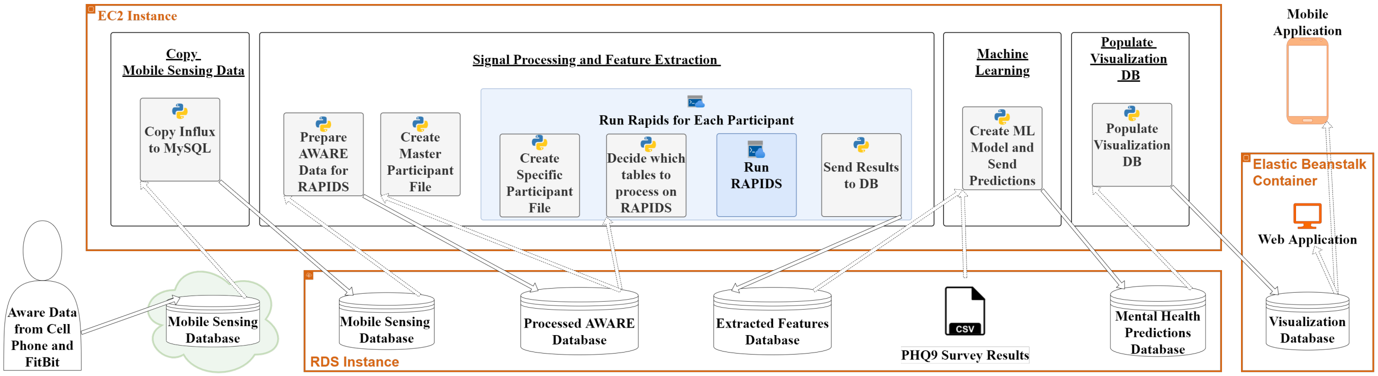


Fig. 1. Data Pipeline Implementation

using the REDCap API. Due to the lack of automation of this step, it has not been included in any of the figures. We matched the PHQ-9 scores and extracted features for each participants through a script on a weekly basis. Any extracted feature or survey result that did not have a match were dropped.

C. Machine Learning

With the data processing steps completed, the data was passed through a machine learning script to output depression predictions. RAPIDS outputs a large number of columns, many of which do not contain any data if patients turn off various sensors on their phones. In order to preserve useful information, all empty columns and rows, rows or columns with 50% or more missing data, and redundant columns with a correlation of over 0.95 were dropped to reduce multicollinearity. Median imputation was performed in the remaining dataset as many of the features were very skewed. The numerical columns were then scaled between 0 and 1 before training began.

To choose the final algorithms and validation strategy for implementation, we ran an exploratory analysis on data from 40 adolescents using different regression algorithms as listed in Table I using root mean square error (RMSE) as the performance measure.We chose XGBoostRF for the implementation as it provided the lowest RMSE. XGBoostRF stands for Extreme Gradient Boosting Random Forests and combines gradient boosting with an ensemble of decision trees to output a numerical prediction for depression score.

TABLE I
AVERAGE RMSE OF VARIOUS MACHINE LEARNING MODELS ON PATIENT DATA

Algorithm	Average RMSE
XGBoost	5.8
XGBoostRF	**5.3**
Random Forests	5.6
ExtraTrees	5.6
Gradient Boosting	5.8
Adaboost	5.4
Light GBM	6.0
Catboost	5.7

1) Validation Strategy and Performance Measures: Our approach used data aggregated on a weekly basis. Following the method in [24], we evaluated four validation strategies, two based on individual patient data and two based on the population data. The individual models include leave-one-patient-one-week-out (Lopowo) and leave-accumulated-weeks-out (Lawo). The former uses each patient's data alone to predict that patient's mental state for a given week. The latter uses a patient's data from weeks prior to the current week to predict the patient's mental state for that week. We also evaluated two unified patient model validation approaches namely leave-one-patient-out (Lopo) and leave-one-week-out (Lowo). Lopo uses all other patients' data to predict a given patient's mental state whereas Lowo uses all other weeks of data from all patients to predict a patient's mental state for that week.

To evaluate the automated modeling with continuous new data, we ran regression and classification experiments with the four validation strategies stated above. We first divided the data into nine sections each corresponding to one month of data. Each section added more recent data to the previous section, mimicking the increase in patient data collected by the app over time. The cross validation methods were tested across the nine increasing sections. Table II shows the results of each method using the XGBoostRF regressor and indicates that the Lawo strategy was the best validation approach for the data as it had the lowest RMSE on the final iteration. We therefore, chose to run the machine learning script on the accumulated patients' data on a weekly basis to train a model and to predict the depression score of the current week. The output of this model including the predicted depression score and the important features involved in modeling are stored in another database to be used on the web portal.

As previously mentioned, we also designed a mobile application for patients to monitor their mental health status. The design uses categories of depression (minimal to severe) instead of the actual PHQ-9 score. Therefore, we ran a classification test to choose the algorithm and the validation strategy. We initially converted the PHQ-9 scores into minimal, mild, moderate, moderately severe and severe.

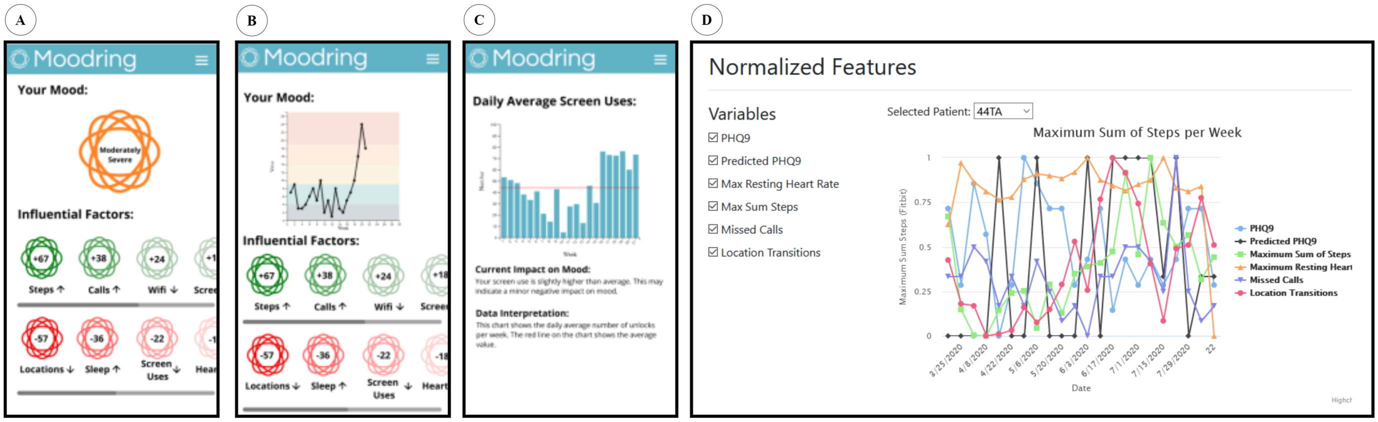


Fig. 2. Moodring App Preview

However, because of the sparsity in some categories and to balance the classes, we merged the categories to create labels "Minimal or Mild" (scores 0-9), "Moderate" (scores 10-14), and "Moderately Severe or Severe" (scores 15-27). We applied the classification version of XGBoostRF to predict the patient depression category. Each cross validation approach was tested using the classifier to see how they performed on iterations with increasing amounts of patient data (Table II). Consistent with the regression approach, the individual patient approaches show higher accuracy across all iterations than the unified patient validation approaches. This indicates that using personalized models for predicting patient PHQ-9 is more accurate than using models that pool data from a group of individuals.

TABLE II
XGBOOSTRF PREDICTION EVALUATION OVER INCREASING DATA

	Lopo		Lowo		Lopowo		Lawo	
Iter.	C Acc	R RMSE	C Acc	R RMSE	C Acc	R RMSE	C Acc	R RMSE
1	11%	3.8	50%	3.4	88%	2.7	86%	2.6
2	38%	4.6	70%	3.6	73%	3.1	77%	2.6
3	44%	5.6	70%	4.2	74%	3.5	74%	3.2
4	47%	6.3	67%	4.6	80%	3.5	80%	3.3
5	47%	5.9	64%	4.7	75%	3.5	75%	3.3
6	41%	5.7	62%	4.8	73%	3.3	73%	3.2
7	40%	5.7	60%	4.8	70%	3.4	71%	3.3
8	41%	5.7	59%	4.8	72%	3.5	73%	3.5
9	42%	5.8	59%	4.7	73%	3.4	73%	3.4

D. Visualization

The final step in the data pipeline was to automatically display and visualize the extracted features and machine learning predictions for further analysis by clinicians and researchers. A final product of the system would involve separate views and usability for members of the primary user classes of adolescents, parents, and care providers. Our current implementation is a dynamic web portal developed primarily for use as a developer and research portal seen in Fig. 2 D, but this implementation is also useful for clinicians to interpret the machine learning outputs. The front end is a Django-based application that is hosted on an Amazon Elastic Beanstalk system. The Django application accesses a database holding selected features and user PHQ-9 predictions and communicates these to the user by dynamically creating pages displaying line plots of normalized RAPIDS data for each feature used in the machine learning script separated by sensor. The features displayed are read from a generic configuration file that is updated based on the results of the machine learning The user has the ability to turn on and off each features in order to investigate specific trends as well as overlay the predicted and ground truth PHQ-9. In addition to the functional web portal, we also developed wireframes for a phone based application for adolescent users shown in Fig. 2 A, B and C. The design was refined through interviews with users and non-users. The mobile application utilizes feature importance from the machine learning algorithm and the classified PHQ-9 scores to display the predicted depression category as well behavioral features that contributed most to the prediction of that category. This gives users insights into what behavioral factors correspond to their depressive symptoms. The feature importance, direction, and intensity are presented using color, shade, and numerical labels.

IV. CONCLUSION AND FUTURE WORK

We presented an automated machine learning pipeline for continuous processing and analysis of mobile health data and communicating the results back to the stakeholders including clinicians, patients, and researchers. We implemented the pipeline for the real-time assessment of mental health in adolescents. Although we only used depression as a case for evaluation, the design and implementation can be extended to other health and behavioral outcomes with more biobehavioral data e.g., heart rate variability and skin temperature.

Future steps include allowing for simultaneous prediction of multiple outcomes, e.g., anxiety, stress, depression, and sleep quality. The pipeline can further be expanded to allow users to choose which combination of mental state forecasts they want to receive.

We plan to add functionality to allow clinicians and researchers to specify the analysis settings including selection of the machine learning algorithm, identifying the prediction outcome, specifying the length of data to be used in the analysis, etc. This step will add more flexibility and value to the automated pipeline and help both clinicians and researchers draw insights in a ongoing basis.

We also plan on more in depth HCI research on the stakeholder interactions with the application to further identify key functionalities of the system. For example, an essential functionality for clinicians might be to give them an overview of patients at risk. This step will help the full implementation of the mobile application and web portal with added functionality and visualizations for all stakeholders.

References

[1] R. F. Dickerson, E. I. Gorlin, and J. A. Stankovic, "Empath: A continuous remote emotional health monitoring system for depressive illness," in *Proceedings of the 2nd Conference on Wireless Health*, WH '11, (New York, NY, USA), Association for Computing Machinery, 2011.

[2] D. C. Mohr, K. N. Tomasino, E. G. Lattie, H. L. Palac, M. J. Kwasny, K. Weingardt, C. J. Karr, S. M. Kaiser, R. C. Rossom, L. R. Bardsley, L. Caccamo, C. Stiles-Shields, and S. M. Schueller, "Intellicare: An eclectic, skills-based app suite for the treatment of depression and anxiety," *J Med Internet Res*, vol. 19, p. e10, Jan 2017.

[3] A. Ghandeharioun, S. Fedor, L. Sangermano, D. Ionescu, J. Alpert, C. Dale, D. Sontag, and R. Picard, "Objective assessment of depressive symptoms with machine learning and wearable sensors data," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 325–332, 2017.

[4] A. Pratap, D. C. Atkins, B. N. Renn, M. J. Tanana, S. D. Mooney, J. A. Anguera, and P. A. Areán, "The accuracy of passive phone sensors in predicting daily mood," *Depression and Anxiety*, vol. 36, no. 1, pp. 72–81, 2019.

[5] A. A. Farhan, C. Yue, R. Morillo, S. Ware, J. Lu, J. Bi, J. Kamath, A. Russell, A. Bamis, and B. Wang, "Behavior vs. introspection: refining prediction of clinical depression via smartphone sensing data," in *2016 IEEE Wireless Health (WH)*, pp. 1–8, 2016.

[6] D. Ferreira, V. Kostakos, and A. K. Dey, "Aware: Mobile context instrumentation framework," *Frontiers in ICT*, vol. 2, p. 6, 2015.

[7] D. C. Mohr, K. N. Tomasino, E. G. Lattie, H. L. Palac, M. J. Kwasny, K. Weingardt, C. J. Karr, S. M. Kaiser, R. C. Rossom, L. R. Bardsley, L. Caccamo, C. Stiles-Shields, and S. M. Schueller, "Intellicare: An eclectic, skills-based app suite for the treatment of depression and anxiety," *J Med Internet Res*, vol. 19, p. e10, Jan 2017.

[8] J. Torous, H. Wisniewski, B. Bird, E. Carpenter, G. David, E. Elejalde, D. Fulford, S. Guimond, R. Hays, P. Henson, *et al.*, "Creating a digital health smartphone app and digital phenotyping platform for mental health and diverse healthcare needs: an interdisciplinary and collaborative approach," *Journal of Technology in Behavioral Science*, vol. 4, no. 2, pp. 73–85, 2019.

[9] Y. Ranjan, Z. Rashid, C. Stewart, P. Conde, M. Begale, D. Verbeeck, S. Boettcher, R. Dobson, and A. Folarin, "Radar-base: Open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices," *JMIR Mhealth Uhealth*, vol. 7, p. e11734, Aug 2019.

[10] H. Xiong, Y. Huang, L. E. Barnes, and M. S. Gerber, "Sensus: A cross-platform, general-purpose system for mobile crowdsensing in human-subject studies," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '16, (New York, NY, USA), p. 415–426, Association for Computing Machinery, 2016.

[11] J. E. Bardram, "The carp mobile sensing framework – a cross-platform, reactive, programming framework and runtime environment for digital phenotyping," 2020.

[12] J. Torous, M. V. Kiang, J. Lorme, and J.-P. Onnela, "New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research," *JMIR Mental Health*, vol. 3, p. e16, May 2016.

[13] J. Vega, "Monitoring parkinson's disease progression using behavioural inferences, mobile devices and web technologies," in *Proceedings of the 25th International Conference Companion on World Wide Web*, WWW '16 Companion, (Republic and Canton of Geneva, CHE), p. 323–327, International World Wide Web Conferences Steering Committee, 2016.

[14] I. Moshe, Y. Terhorst, K. Opoku Asare, L. B. Sander, D. Ferreira, H. Baumeister, D. C. Mohr, and L. Pulkki-Råback, "Predicting symptoms of depression and anxiety using smartphone and wearable data," *Frontiers in Psychiatry*, vol. 12, p. 43, 2021.

[15] S. Bae, T. Chung, D. Ferreira, A. K. Dey, and B. Suffoletto, "Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions," *Addictive Behaviors*, vol. 83, pp. 42–47, 2018. Ambulatory Assessment of Addictive Disorders.

[16] A. Doryab, D. K. Villalba, P. Chikersal, J. M. Dutcher, M. Tumminia, X. Liu, S. Cohen, K. Creswell, J. Mankoff, J. D. Creswell, and A. K. Dey, "Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: Statistical analysis, data mining and machine learning of smartphone and fitbit data," *JMIR Mhealth Uhealth*, vol. 7, p. e13209, Jul 2019.

[17] P. Chikersal, A. Doryab, M. Tumminia, D. K. Villalba, J. M. Dutcher, X. Liu, S. Cohen, K. G. Creswell, J. Mankoff, J. D. Creswell, M. Goel, and A. K. Dey, "Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: A machine learning approach with robust feature selection," *ACM Trans. Comput.-Hum. Interact.*, vol. 28, Jan. 2021.

[18] S. Shiffman, A. A. Stone, and M. R. Hufford, "Ecological momentary assessment," *Annu. Rev. Clin. Psychol.*, vol. 4, pp. 1–32, 2008.

[19] J. Vega, M. Li, K. Aguillera, N. Goel, E. Joshi, K. C. Durica, A. R. Kunta, and C. A. Low, "Rapids: Reproducible analysis pipeline for data streams collected with mobile devices," 08 2020.

[20] A. Doryab, P. Chikarsel, X. Liu, and A. K. Dey, "Extraction of behavioral features from smartphone and wearable data," 2019.

[21] B. Brent, B. Lu, J. Kim, and J. P. Dunn, "Biosignal compression toolbox for digital biomarker discovery," *Sensors (14248220)*, vol. 21, no. 2, pp. 526–527, 2021.

[22] X. Wang, N. Vouk, C. Heaukulani, T. Buddhika, W. Martanto, J. Lee, and R. J. Morris, "Hopes: An integrative digital phenotyping platform for data collection, monitoring, and machine learning," *J Med Internet Res*, vol. 23, p. e23984, Mar 2021.

[23] V. P. Cornet and R. J. Holden, "Systematic review of smartphone-based passive sensing for health and wellbeing," *Journal of Biomedical Informatics*, vol. 77, pp. 120–132, 2018.

[24] A. Doryab, M. Frost, M. Faurholt-Jepsen, L. V. Kessing, and J. E. Bardram, "Impact factor analysis: combining prediction with parameter ranking to reveal the impact of behavior on health outcome," *Personal and Ubiquitous Computing*, vol. 19, no. 2, pp. 355–365, 2015.

[25] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, and F. Hutter, "Efficient and robust automated machine learning," in *Advances in Neural Information Processing Systems* (C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, eds.), vol. 28, Curran Associates, Inc., 2015.

[26] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Auto-weka: Combined selection and hyperparameter optimization of classification algorithms," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '13, (New York, NY, USA), p. 847–855, Association for Computing Machinery, 2013.

[27] M. Feurer, A. Klein, K. Eggensperger, J. T. Springenberg, M. Blum, and F. Hutter, "Auto-sklearn: efficient and robust automated machine learning," *Automated Machine Learning*, pp. 113–134, 2018.

[28] S. Abdullah, M. Matthews, E. Frank, G. Doherty, G. Gay, and T. Choudhury, "Automatic detection of social rhythms in bipolar disorder," *Journal of the American Medical Informatics Association*, vol. 23, pp. 538–543, 03 2016.

[29] M. Frost, A. Doryab, M. Faurholt-Jepsen, L. V. Kessing, and J. E. Bardram, "Supporting disease insight through data analysis: refinements of the monarca self-assessment system," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pp. 133–142, 2013.

[30] K. Kroenke, R. L. Spitzer, and J. B. Williams, "The phq-9: validity of a brief depression severity measure," *Journal of general internal medicine*, vol. 16, no. 9, pp. 606–613, 2001.

Confronting the Duality of Using Potentially Harmful Technology as Host to Beneficial Technology

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Joseph Keogh
Spring, 2021

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature Joseph Keogh Date May 11, 2021
Joseph Keogh

Approved ______________________________ Date __________
Richard Jacques, Department of Engineering and Societ

Confronting the Duality of Using Potentially Harmful Technology as Host to Beneficial Technology

Motivation

Adolescent depression is a public health problem, with 3.2 million adolescents (aged 12–17), which represents roughly 13 percent of the population having a major depressive episode in 2017 in the United States alone (NIMH, 2020, Prevalence of Major Depressive Episode Among Adolescents section). Only 40 percent (1.2 million) of adolescents with major depression in the United States were receiving treatment (NIMH, 2020, Treatment of Major Depressive Episode Among Adolescents section), leaving 1.8 million adolescents without treatment. Thus, major depressive episodes remain untreated in many cases where the individual is an adolescent. Wisdom et al. (2006) also discuss why teens are not willing to seek treatment. A consequence of this immaturity is an unwillingness to confide in parents and guardians. As a result of this issue, care becomes more difficult.

We have, through interviewing mental healthcare providers, discovered that the largest discrepancy in providing mental health care to adolescents is simply knowing who needs help and when to give it to them. Health providers claim to often see young patients with severe depression where the parents had no idea the teen was suffering (A. Radovic MD, personal communication, October 17, 2020).

The project will attempt to assist in the diagnosis and monitoring of adolescents with depression. This technology would allow teens, parents, and caretakers to monitor the adolescent's depression. Allowing intervention when the adolescent is deteriorating in tandem with treatment being relaxed when the adolescent is improving. Facilitating greater efficiency in the care of the adolescent's mental health.

By using existing sensing technology inside smartphones along with novel technology, we will develop that includes machine learning principles, we will monitor and forecast the mental health state of the adolescent. This monitoring and forecasting will help all parties learn about and be more involved in the mental health state of the adolescent, hopefully providing a springboard for a mental healthcare increase.

It is recognized that a duality is created through this new technology. This duality is the use of smartphones, potentially harmful to the teen's mental health, to host a technology that is potentially beneficial to mental health. During the project the duality will be considered and examined to ethically validate our technology to ensure it will provide a net positive to the current mental health crisis.

The Moodring project is sponsored by the United States National Institute of Health, and is a collaboration between the University of Pittsburgh, the University of Virginia, and a Pittsburgh-area software company called NuRelm. At the University of Virginia, Afsaneh Doryab holds the grant from the National Institute of Health. Throughout the process of developing Moodring, the team will write a technical paper on the project. The team will submit the technical paper to the Systems and Information Engineering Design Symposium conference in April.

Introduction

As previously mentioned, there is a significant mental health crisis in America that is affecting adolescents. To help combat this crisis, we are attempting to develop technology that will aid in monitoring and forecasting the mental health status of adolescents using passive sensing and machine learning. This technology will be deployed via a downloadable application on the user's cell phone.

There has been research that shows that cell phones are contributing to the mental health degradation of teenagers. Alhassan et al. have shown for every unit increase in cell phone addiction (Smart-phone Addiction scale of 0–100) there is a one unit increase in the user's depression (Beck Depression scale 0–40). With 64 percent of American's being slight addicts of cell phones (Alhassan, 2020), we can see that cell phones can have inherent danger to the user's mental health. Another study by Chen Y. specifically looked at college-aged students and found a significant pattern between cell phone addiction and depression amongst the students (Chen Y, 2004).

The specific problem that will be addressed is looking at is it ethical to provide a possible solution to a harmful problem by providing that solution through a potentially harmful medium? Our specific problem being is it ethical to use a cell phone, which we know can deteriorate the mental health of adolescents, to possibly improve a user's mental health?

This duality is analogous to providing medication that could have potentially crippling side effects. While prescription drugs can combat diseases and ailments, there is a 20 percent chance of a serious negative reaction, leading prescription drug side effects to be the 4th leading cause of death, about 328,000 deaths in America and Europe each year (Light, 2014). With prescription medication side effects being very serious, we must take our own research just as serious as they are both dealing with the health of the users.

If we do not attempt to confront this duality, we could be creating potentially life-threatening technology. This threat coming from the risk of suicide as depression worsens due to the increased cell phone use. While it is not guaranteed that the technology we produce would be harmful, we must do due diligence to ensure that our technology is safe. While we may not come to a solid answer on whether our technology is effective and safe, which should be done through verified trials, we can at least begin the process and conversation into the safety of our technology.

If we can effectively confront this duality, even if we fail at creating the technology, we will be able to benefit the research and medical communities. With the duality confronted, we will know more about how technology can be used to benefit the adolescent and their mental health while mitigating the risk that the actual smartphones provide. This information will be useful to future designers for the same or similar technology that we are attempting to create. They will be able to create their own technology knowing what boundaries they can and cannot cross, and what role their technology will play in the social context of the user. This information will give developers the confidence and ability to create a truly useful technology. Of course, if we are successful in our design, we will have this same confidence. Unfortunately, both research efforts are occurring concurrently, so it is possible that our technical deliverable will not align completely with our STS research.

Literature Review

While this specific topic, confronting putting potential beneficial mental health technology in a potentially harmful medium, there has been research conducted in related fields: medicinal side effects and the cost benefit of using medicines, and efficacy of mental health medicine regulation applications. We will review these two main topics below.

Existing Health App Review

Most medication applications focus on medication adherence (Ankem, 2019). These applications are a much-needed aid in today's world considering nearly 75% of adults are non-adherent to their prescription medications (Ankem, 2019). Medication adherence refers to the patient taking their medication dose when they are prescribed, considering the time of day, frequency, and dosage amount. Similar to medication adherence, medication safety is also very important. Medication safety is defined in this context as the prescriber of the medication not making an error. An error could be improper dosage amount, improper timing, or combining medications together that could be dangerous.

A qualitative study was conducted on the ScopiaRx medication safety app that was developed the Center for Applied Informatics (CAI) in 2013. The application is designed to be used by patients, caretakers, and medical professionals. Patients and caretakers use the app to increase the medication adherence of the patient. Medical professionals can use the application during the medication prescription phase. Physicians and pharmacists use the database of medications that is provided in the app to better understand the dosage, timing, and drug combinations that they are considering for the patient. Allowing for safer dosage and drug combinations to be prescribed to the patient.

The study conducted was a usability test on the ScopiaRx application. The goals of the study were to determine if RN-BSN (Registered Nurse-Bachelor Science in Nursing) students can accurately guess the meaning of app components and perform tasks in a short time period with the app. 18 RN-BSN students participated in the study.

The main findings of the study was that if medication safety apps are to be used in the medical world, they must allow an intuitive interface, match the real world in which healthcare workers operate, and facilitate the management of information through optimal information architecture. The study also found that a younger generation is more likely to engage with medication applications.

Side effects of mental health medications

On the same health topic as our discussion, a study was conducted into the side effects of mental health medications in Nepal (Marasine, 2020). This study parallels our conversation as both mediums (medications or our mental health application) are attempting to better the users' (or patients') mental health. 174 patients were considered during the study.

Antidepressant medication is often deemed the best treatment option for depression (Marasine, 2020). Examples of antidepressant medications are: serotonin reuptake inhibitors, serotonin and noradrenaline reuptake inhibitors, and selective serotonin-noradrenaline reuptake inhibitors. The authors note that when prescribing antidepressants, the medical professional should also take other factors into consideration such as adverse effect profiles, cost, safety profile, history of prior medication treatment, and patient preference. Marasine et al. mention that antidepressant drug side effects are the significant determinant of medication nonadherence.

The study aimed to evaluate antidepressant side effects, medication adherence and factors contributing to depression.

Marasine et al. report that ~75% of patients encountered side effects while taking antidepressants. The breakdown of side effects are as follows: insomnia and anxiety 17%, dry mouth 10% and weight gain 10%, with ~70% of the patients suffered moderate (56%) to severe (12%) side effects.

Table inspired by Marasine et al.
“Severity of side effects and probability of adverse drug reactions (n=129)”

Variables		n (% of total study)
	Mild	9 (5)
Severity	Moderate	98 (56)
	Severe	22 (12)
Probability	Probable	108 (83)
	Possible	21 (16)

Cost of bad drugs

The cost to individuals of Treatment-related problems (TRPs) have a significant impact on the health and economic status of Americans (Al-Qudah, 2020). For example, the costs associated with adverse drug events doubled between 1995 and 2000 ($76.6B to $177.4B).

A retrospective study was conducted to analyze the cost of medication errors in individuals admitted to the hospital (McCarthy, 2017). Four hundred sixteen (416) hospitalization cases were evaluated. McCarthy et al. found that patients who experienced adverse drug reactions had an average cost of $19k per visit while those that did not experience adverse drug reactions had an average cost of $17k. The cost difference is roughly 10% in those patients that had adverse drug reactions. Adverse drug reactions are defined as “an injury resulting from the use of a drug” as defined by McCarthy et al.

Adverse drug reactions result in more than 770,000 injuries or deaths in United States hospitals each year (Slight, 2018). Slight et al. claim that most of these reactions are preventable and could be reduced by hospitals making changes to their drug ordering systems. This issue is similar to what the ScopiaRx app was attempting to solve. With ScopiaRx operating on a small individual scale. Slight et al. recommend that computerized provider order entry (CPOE) and clinical decision support (CDS) both have the potential to improve the drug ordering process. These systems alert the medical staff when a potential hazardous drug to drug interaction has been identified. The study found that an estimated 5.5 million drug alerts (7%) were incorrectly overridden by the healthcare provider resulting in ~200,000 adverse drug reactions costing a total of $870M-$1760M to the patients. This averages $4350 per adverse drug reaction due to an incorrect decision made by healthcare professionals. (Slight, 2018).

Double Edged Sword

Shaw et al. claim that more than 100,00 deaths occur each year in the United States due to adverse drug reactions. To prevent adverse drug reactions, pharmaceutical companies perform safety trials. One issue with these trials, claim Shaw et al., is that most these trials are conducted on individuals with European decent. The result is a drug that may be safe for that class of individuals, but may not be for those with a more diverse genetic lineage. A focus group was formed comprised of Alaskan Natives (an example of an underrepresented group in pharmaceutical studies of safety). The goal of the focus group was to inquire about if the individuals would be interested in the Alaskan Native community becoming more represented in the pharmaceutical community. The general consensus of the group was that they are open to exploring more western medicinal practices, but recognize that there are risks involved. The group concluded that western medicine could bring a safer, more efficient diagnosis and treatment of chronic and life-threatening health conditions. However, the group was concerned about the potential physical and cultural harm that could be brought with the possible adaptation. The authors finally offer the following advice, "the participants' recognition of the dual potential of pharmacogenetics tells a cautionary tale for anyone working in communities where medical practice and research has perpetrated harm or cultivated distrust. Pharmacogenetic research must be undertaken with respect for these histories, restoration of trust, and responsiveness to the communities in which we work."

Discussion

Below we will discuss the measures to address the duality of using a potentially harmful medium to host a beneficial technology.

Cost Benefit Analysis

To best understand whether, on average in the long, a decision is worthwhile or not a technique called 'Cost Benefit Analysis' is used. This technique combines the value of an outcome, with the probability of that outcome occurring. Performing this analysis across all possible outcomes gives you the 'expected value' of the decision.

$$\sum(Probability_i * Value_i) = Expected\ Value$$

$$\sum(P_i * V_i) = E$$

For cost benefit analysis to be most meaningful, confidence in both the probability and value of the outcome are needed. Taking information from the literature above, we can provide estimates for the expected value of a patient taking antidepressant medication. We will assume that whether there is an adverse drug reaction (ads) or not, the medication will alleviate the depression of the user. Further research should be conducted into the probability of success for antidepressant medications.

$$P_{ads} * V_{ads} + P_{no-ads} * V_{no-ads} = E_{medication}$$

$$70\% * {}^{1}(-\$3175) + 30\% * (\$0) = E_{medication}$$

$$-\$2222.50 \approx -\$2000 = E_{medication}$$

It is recognized that all probabilities and values used above are estimates based solely on prior literature and that only financial gain and loss are taken into consideration. Further research is recommended into putting a numerical value on the physical toll of an adverse drug reaction.

We can now compare the patient taking an antidepressant medication to them using Moodring. We cannot estimate the values or probabilities of Moodring, however we can define them.

$$P_{harmful} * V_{harmful} + P_{noHarm} * V_{moodring} = E_{moodring}$$

---

[1] $3175 is derived from the literature review ($4350+$2000)/2 to find the average cost of an adverse drug reaction

In the first equation above, we find the expected value of using a cell phone. This will likely depend on the user's own personality, bias, addiction, and specific mental health, very much related to how the study conducted by Marasine examined how different factors contributed to depression. Followed by aggregating the expected value of the user using the cell phone, with the added value of using Moodring.

We can simplify the equations and relate the expected value of using medication to the expected value of using Moodring. If the expected value of using Moodring is greater than the expected value of using traditional medications, the user will have another tool in their decision making.

$$P_{harmful} * V_{harmful} + \left(1 - P_{harmful}\right) * V_{moodring} = E_{moodring} > E_{medication}$$
$$P_{harmful} * V_{harmful} + \left(1 - P_{harmful}\right) * V_{moodring} > E_{medication}$$
$$P_{harmful} > \frac{E_{medication} - V_{moodring}}{V_{harmful} - V_{moodring}}$$
$$P_{harmful} > \frac{-\$2000 - V_{moodring}}{V_{harmful} - V_{moodring}}$$

From the final equation above, we can now give a definitive answer into whether it is potentially beneficial for a user to use Moodring. We are not claiming that this analysis should drive the patients decision 100%, we are only providing it as another source of information for decision making. Again, the above equations only consider the financial aspect of the decision making however, the logic behind the symbolic answers is very sound. Once the expected value of taking medication, the value of Moodring, and the potential harm from using a cell phone are properly defined, this cost benefit analysis will be invaluable for those struggling with depression in their pursuit of relief.

Ensuring Diversity

A lesson to be learned from the Alaskan Native study is the consideration of diversity in our product development. Before full implementation proper consideration should be given to the feasibility of the product for a range of users. Our product, and any of the sort, should be tested with user of different: races, age, sex, gender, experience, health conditions, and many others. This will help improve the overall operability of the product, as well as the trust in the product from the public.

Communicating the Situation to Users

The final advise to technology creators of this nature is very straightforward and powerful. The best way to ensure you are not violating ethics, overstepping, or overpromising your technologies capabilities is to be transparent with your customers and product users. We have shown through our cost benefit analysis that is not possible to fully quantify the expected value of using a potential beneficial technology that is hosted in a potentially dangerous medium for all users. We have shown however, that an individual, given the right information, can make this decision. It is our responsibility to provide the user with all the information we can so they can decide on whether to take the risk in using our product.

Conclusion

We have shown through our literature review, cost benefit analysis, and other discussions, that it is possible to provide a potentially beneficial technology in a potentially harmful medium as long as we are properly categorizing and setting realistic expectations for our users. This directly correlates to our project of putting mental health assistance technology in cell phones. If we can properly communicate the possible benefits and downsides of our technology, we can empower the user to make an informed decision if using this technology is right for them.

The goal in all health sciences should be to solely benefit society. Fortunately or unfortunately the responsibility of ensuring safe technology be put into the public does not fall on the engineers, scientists, or developers of that technology. We must do everything we can to use our knowledge of the technology to inform the public on its true value to them. This will allow for more transparency and eventually a healthier population. With health being the true priority of the medical industry man-kind will be propelled into a new level of being.

References

Alhassan, A. A., Alqadhib, E. M., Taha, N. W., Alahmari, R. A., Salam, M., & Almutairi, A. F. (2018). The relationship between addiction to smartphone usage and depression among adults: A cross sectional study. *BMC Psychiatry*, *18*. https://doi.org/10.1186/s12888-018-1745-4

Al-Qudah, R. A., Al-Badriyeh, D., Al-Ali, F. M., Altawalbeh, S. M., & Basheti, I. A. (2020). Cost-benefit analysis of clinical pharmacist intervention in preventing adverse drug events in the general chronic diseases outpatients. *Journal of Evaluation in Clinical Practice*, *26*(1), 115–124. https://doi.org/10.1111/jep.13209

Ankem, K., Cho, S., & Simpson, D. (2019). Nurses' perceptions and problems in the usability of a medication safety app. *Informatics for Health & Social Care*, *44*(1), 48–69. https://doi.org/10.1080/17538157.2017.1364249

Chen, Y.-F. (2004, October). *The relationship of mobile phone use to addiction and depression amongst American college students*. 한국방송학회 세미나 및 보고서. https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE01121669

Light, D. (2020, June 27). *New Prescription Drugs: A Major Health Risk With Few Offsetting Advantages*. Harvard Center for Ethics. https://ethics.harvard.edu/blog/new-prescription-drugs-major-health-risk-few-offsetting-advantages

Mao, J., Li, T., Fan, D., Zhou, H., Feng, J., Liu, L., Zhang, C., & Wang, X. (2020). Abnormal expression of rno_circRNA_014900 and rno_circRNA_005442 induced by ketamine in the rat hippocampus. *BMC Psychiatry*, *20*(1), 1. https://doi.org/10.1186/s12888-019-2374-2

Marasine, N. R., Sankhi, S., Lamichhane, R., Marasini, N. R., & Dangi, N. B. (2020). Self-Reported Antidepressant Drug Side Effects, Medication Adherence, and Its Associated Factors among

Patients Diagnosed with Depression at the Psychiatric Hospital of Nepal. *Depression Research & Treatment*, 1–7. https://doi.org/10.1155/2020/7024275

McCarthy Jr., B. C., Tuiskula, K. A., Driscoll, T. P., & Davis, A. M. (2017). Medication errors resulting in harm: Using chargemaster data to determine association with cost of hospitalization and length of stay. *American Journal of Health-System Pharmacy*, *74*, 102–107. https://doi.org/10.2146/ajhp160848

NIMH » Major Depression. (n.d.). Retrieved November 10, 2020, from https://www.nimh.nih.gov/health/statistics/major-depression.shtml

Shaw, J. L., Robinson, R., Starks, H., Burke, W., & Dillard, D. A. (2013). Risk, Reward, and the Double-Edged Sword: Perspectives on Pharmacogenetic Research and Clinical Testing Among Alaska Native People. *American Journal of Public Health*, *103*(12), 2220–2225. https://doi.org/10.2105/AJPH.2013.301596

Slight, S. P., Seger, D. L., Franz, C., Wong, A., & Bates, D. W. (2018). The national cost of adverse drug events resulting from inappropriate medication-related alert overrides in the United States. *Journal of the American Medical Informatics Association : JAMIA*, *25*(9), 1183–1188. https://doi.org/10.1093/jamia/ocy066

Tandon, R., Lenderking, W. R., Weiss, C., Shalhoub, H., Barbosa, C. D., Chen, J., Greene, M., Meehan, S. R., Duvold, L. B., Arango, C., Agid, O., & Castle, D. (2020). The impact on functioning of second-generation antipsychotic medication side effects for patients with schizophrenia: A worldwide, cross-sectional, web-based survey. *Annals of General Psychiatry*, *19*(1), 1–11. https://doi.org/10.1186/s12991-020-00292-5

Wartella, E., Rideout, V., Montague, H., Beaudoin-Ryan, L., & Lauricella, A. (2016). Teens, Health and Technology: A National Survey. *Media and Communication*, *4*(3), 13–23. https://doi.org/10.17645/mac.v4i3.515

Wisdom, J. P., Clarke, G. N., & Green, C. A. (2006). What Teens Want: Barriers to Seeking Care for Depression. *Administration and Policy in Mental Health*, *33*(2), 133–145. https://doi.org/10.1007/s10488-006-0036-4

A Automated Machine Learning Pipeline for the Monitoring and Forecasting of Mobile Health Data
(Technical Paper)

Confronting the Duality of Using Potentially Harmful Technology as Host to Beneficial Technology

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Joseph Keogh
Spring, 2021

Technical Project Team Members
Technical Project Team Members
Ana Bonaquist
Meredith Grehan
Owen Haines
Neil Singh

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature Joseph Keogh Date May 11, 2021
Joseph Keogh

Approved ____ Date __________
Afsaneh Doryab, Department of Systems and Information Engineering

Approved ____ Date __________
Richard Jacques, Department of Engineering and Society

Introduction

Adolescent depression is a public health problem, with 3.2 million adolescents (aged 12-17), which represents roughly 13 percent of the population having a major depressive episode in 2017 in the United States alone (NIMH, 2020, Prevalence of Major Depressive Episode Among Adolescents section). Only 40 percent (1.2 million) of adolescents with major depression in the United States were receiving treatment (NIMH, 2020, Treatment of Major Depressive Episode Among Adolescents section), leaving 1.8 million adolescents without treatment. Thus, major depressive episodes remain untreated in many cases where the individual is an adolescent. Wisdom et al. (2006) discusses also why teens may not be willing to seek out treatment. A consequence of this immaturity is an unwillingness to confide in parents and guardians. As a result of this issue, care becomes more difficult.

We have, through interviewing mental health care providers, discovered that the largest discrepancy in providing mental health care to adolescents is simply knowing when and who to give help to. Health providers claim to often do they see patients with severe depression where the parents had no idea the teen was suffering (A. Radovic MD, personal communication, October 17, 2020).

The project will attempt to assist in the diagnosis and monitoring of adolescents with depression. This technology would allow teens, parents, and caretakers to monitor the adolescent's depression. This would allow intervention when the adolescent is deteriorating in tandem with treatment being relaxed when the adolescent is improving. Facilitating greater efficiency in the care of the adolescent's mental health.

By utilizing existing sensing technology inside of smart phones along with novel technology we will develop that includes machine learning principles, we will monitor and forecast the mental health state of the adolescent. This monitoring and forecasting will help all parties learn about and be more involved in the mental health state of the adolescent, hopefully providing a springboard for a mental health care increase.

It is recognized that a duality is created through this new technology. This duality being the use of smart phones which are potentially harmful to the teen's mental health. During the project this will be considered and examined to ethically validate our technology to ensure it will provide a net positive to the current mental health crisis.

The Moodring project is sponsored by the United States National Institute of Health, and is a collaboration between the University of Pittsburgh, the University of Virginia, and a Pittsburgh-area software company called NuRelm. At the University of Virginia, Afsaneh Doryab holds the grant from the National Institute of Health. Throughout the process of developing Moodring, the team will write a technical paper on the project. The team will submit the technical paper to the Systems and Information Engineering Design Symposium conference in April.

Technical Topic: Utilizing Passive Sensing to Monitor and Forecast Depression in Adolescents

According to Harari et al. (2016), "In existing procedures for collecting data on behavior, researchers typically ask participants to estimate the frequency or duration of past or typical behaviors" (p. 839). This approach to patient self-monitoring, accomplished via survey data, serves as a major tool for diagnosing and tracking depression. However, self-monitoring for depression symptoms presents several shortcomings. Patients with depression may be less likely to fill out surveys, since consistently recognized symptoms have included fatigue and lack of motivation (Substance Abuse and Mental Health Services Administration, 2016, Table 29). Likewise, patients not actively experiencing symptoms may not feel that self-monitoring is necessary. These issues are further exacerbated in adolescent patients who might be reticent to share survey information with parents or healthcare providers.

Self-monitoring of depressive symptoms relies on the input and participation of the user. This can create problems due to the mental state of that user. Current mobile applications for adolescent mental health care utilize self-monitoring and therefore cannot provide optimal user experience nor accuracy regarding mental state. The technology we are developing, Moodring, would avoid self-monitoring issues by not requiring active input from patients to track and predict mental health states, facilitated by passive sensing and machine learning. According to Cornet and Holden (2017), passive sensing describes "...the capture of data about a person without extra effort on their part" (p. 120). Luxton (2016) explains machine learning as "…a core branch of AI [artificial intelligence] that aims to give computers the ability to learn without being explicitly programmed" (p. 3).

Monitoring adolescent depression with passive sensing builds on previous approaches to combat depression. For example, in 2013, Frost et al. used passive sensing in an application called Monarca to successfully estimate manic and depressive states in adults with bipolar disorder using only sensor data (p. 142). In 2011, Burns et al. created an application called "Mobilyze!," which successfully used machine learning models to predict patient mood, emotion, and motivational state based on a variety of sensor values from smartphones (p. 1). Currently, an application called Monsenso also monitors patients with their smartphone sensors and relays the data to both patients and health care providers (Monsenso). We plan to extend these approaches to adolescents with depression, who may be ideal candidates due to their high levels of smartphone use. Current work in the field of adolescent mental health mobile computing is surveyed by Grist, Porter, and Stallard (2017).

The team will use quantitative data from passive sensing coupled with machine learning to identify and forecast depression symptoms and provide suggestions to alleviate symptoms. We will use the AWARE framework, an existing passive sensing application and codebase, to conduct passive sensing, collecting sensor data from smartphones with features such as location, screen time, phone calls, and distance travelled. The AWARE framework is an open source programming framework that "…captures hardware, software, and human-based data from smartphones" (https://awareframework.com). Using this sensor data, the team will engineer other useful features that may indicate signs of depression such as amount of sleep, amount of time spent at home, and frequency of communication with peers.

Our team will create a machine learning model to provide greater insight into the adolescent's mental state based on features most relevant to the patient. Features such as: location, call data, messaging data, pedometer, sleep, and heart rate all measured via the passive sensing technology. The application will relay this prediction, as well as a summary of the relevant features, to the patient in a dashboard. This approach allows for an individualized experience for each adolescent and provides an opportunity to understand how their behaviors contribute to their mental state. If the patient wants to fill out a questionnaire about their mood, they may do so to provide a basis of comparison for the machine learning model, referred to as the ground truth. However, the application will not require any questionnaire data in order to summarize and forecast mental health status for the patient.

Figure 2 shows the flow of information between the application and the patient. The application will continually process sensor data, forecast mental state, and provide visualizations for the patient. Initially, the team will develop the application using sensor data from their own smartphones to provide a proof of concept of a lightweight smartphone application to forecast mental status and summarize sensor data. The team will test the accuracy of the application's machine learning predictions using de-identified adolescent patient data from Afsaneh Doryab's previous studies.

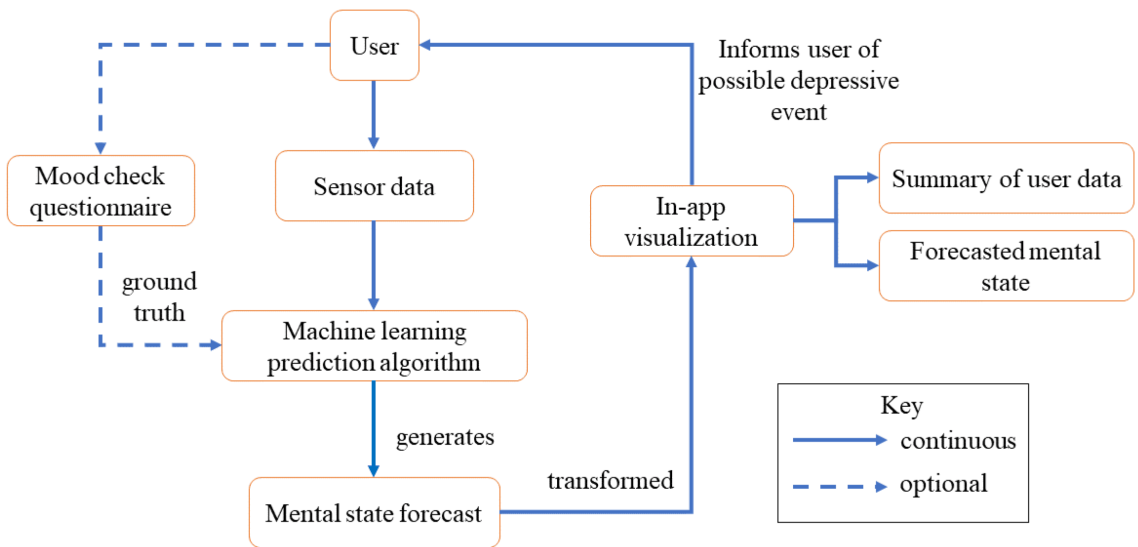


Figure 2. Moodring concept map. This figure lays out how Moodring will collect data and relay information to the user. (Adapted by Anna Bonaquist (2020) from Neil Singh, 2020).

This combination of technology will assist in reaching monitoring and forecasting adolescent depression for several reasons. Passive sensing alleviates the burden of self-monitoring from the adolescent patient. due to the ability to derive information from quantitative measurements from mobile sensors. Previous approaches to self-monitoring allow for user bias to decrease the accuracy and integrity of analysis. Our deliverable will help circumvent self-monitoring flaws by consistently collecting unbiased, objective data through passively monitoring the adolescent. The forecasting methodology for mental health status under the context of depression in adolescents serves as a novel aspect of the deliverable. Currently, technology that monitors mental health exists, but to our knowledge, no other mobile applications can project depressive states of adolescents. Another novel and major improvement over present work is that this modeling will all occur on the device on which the application is installed, a concept known as “on-device machine learning” (Dhar et al., 2020). This will allow for greater privacy over existing practices, which use external storage, such as “the cloud.” This technology has not been developed yet as cell phones are just now gaining the computational ability to complete these tasks.

Our desired outcome is for patients and their caretakers to gain an understanding of the factors affecting their mental state. Specifically, patients should learn what behaviors, like reduced sleep or more time spent at home, contribute to their depression. Another outcome the research provides is an opportunity to facilitate communication between adolescents and their parents and care providers. Adolescents control consent over sharing their data summary and mental health forecast with their parents and care providers. This data sharing would help those who care for the adolescent quickly understand the patient’s mental state and corresponding influential factors.

STS Portion: Confronting the Duality of Using Potentially Harmful Technology as Host to Beneficial Technology

As previously mentioned, there is a significant mental health crisis in America that is affecting adolescents. To help combat this we are attempting to develop technology that will aid in monitoring and forecasting the mental health status of adolescents using passive sensing and machine learning. This technology will be deployed via a downloadable application on the user's cell phone.

There has been research that shows that cell phones are contributing to the mental health degradation of teenagers. Alhassan et al. have shown for every unit increase in cell phone addiction (Smart-phone Addiction scale of 0-100) there is a one unit increase in the user's depression (Beck Depression scale 0-40). With 64 percent of American's being slight addicts of cell phones (Alhassan et al. 2020) (not adjusting for adolescents who may have higher rates) we can see that cell phones may have inherent danger to the mental health of user's. Another study done by Chen Y. looked specifically at college aged students and found a significant pattern between mobile phone addiction and depression amongst the students (Chen Y, 2004).

The specific problem that is going to be addressed is looking at is it ethical to provide a possible solution to a harmful problem by providing that solution through a potentially harmful medium? Our specific problem being: is it ethical to use a cell phone, which we know can deteriorate the mental health of adolescents, to possibly improve a user's mental health?

This duality is analogous to providing medication that could have potentially crippling side effects. While prescription drugs are able to combat diseases and ailments, there is a 20 percent chance of a serious negative reaction, leading prescription drug side effects to be the 4th leading cause of death, about 328,000 deaths in America and Europe each year (Light, 2014). With prescription medication side effects being so serious, we must take our own research just as serious as they are both dealing with the health of the users.

If we do not attempt to confront this duality, we may be creating potentially life-threatening technology. This threat coming from the risk of suicide as depression worsens, the worsening of the depression being caused by increased cell phone use in the user. While it is not guaranteed that the technology we produce would be harmful, we must do due diligence to ensure our technology is safe. While we may not come to a solid answer on whether our technology is effective and safe, which should be done through verified trials, we can at least begin the process and the conversation into the safety of our technology.

If we can effectively confront this duality, even if we fail at creating the technology, we will be able to benefit the research and medical communities. With the duality confronted, we will know more about how technology can be used to benefit the adolescent and their mental health while mitigating the risk that the actual smart phones provide. This information will be useful to future designers for the same or similar technology that we are attempting to create. They will be able to create their own technology knowing what boundaries they can and cannot cross, and what role their technology will play in the social context of the user. This will give them the confidence and ability to create a truly useful technology. Of course if we are successful in our design, we will have this same confidence. Unfortunately, both research efforts are occurring concurrently so it is possible that our technical deliverable will not align completely with our STS research.

Conclusion

The deliverable for this project will be a Minimum Viable Product (MVP) for the client of mobile technology that will monitor and forecast an adolescent's mental health state. This MVP will be in the form of a downloadable mobile application that has cloud technology along side. The user will download the application that will passively monitor them. This data will then be sent up to the cloud where the analysis will occur. Finally, the analysis results will be visualized on a web portal and the downloadable mobile application.

If these deliverables are successfully accomplished, they will contribute to the mental health crisis in a great way. This will be the first product of its kind. No other technology is using passive sensing and machine learning to forecast the mental health of adolescents at this time. With our MVP the client will be able to develop the fully functioning version of the application that will be ready for distribution.

References

Abdullah, S., Matthews, M., Frank, E., Doherty, G., Gay, G., & Choudhury, T. (2016). Automatic detection of social rhythms in bipolar disorder. *Journal of the American Medical Informatics Association*, *23*(3), 538–543. https://doi.org/10.1093/jamia/ocv200

Alhassan, A. A., Alqadhib, E. M., Taha, N. W., Alahmari, R. A., Salam, M., & Almutairi, A. F. (2018). The relationship between addiction to smartphone usage and depression among adults: A cross sectional study. *BMC Psychiatry*, *18*. https://doi.org/10.1186/s12888-018-1745-4

Bardram, J. E., Frost, M., Szántó, K., & Marcu, G. (2012). The MONARCA self-assessment system: A persuasive personal monitoring system for bipolar patients. *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*, 21–30. https://doi.org/10.1145/2110363.2110370

Chen, Y.-F. (2004, October). *The relationship of mobile phone use to addiction and depression amongst American college students*. 한국방송학회 세미나 및 보고서. https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE01121669

Dhar, S., Guo, J., Liu, J., Tripathi, S., Kurup, U., & Shah, M. (2020). On-Device Machine Learning: An Algorithms and Learning Theory Perspective. *ArXiv:1911.00623 [Cs, Stat]*. http://arxiv.org/abs/1911.00623

Doryab, A., Frost, M., Faurholt-Jepsen, M., Kessing, M., & Bardram, J. (2020). Impact factor analysis: Combining prediction with parameter ranking to reveal the impact of behavior on health outcome. *Personal and Ubiquitous Computing*. https://doi.org/10.1007/s00779-014-0826-8

Frost, M., Doryab, A., Faurholt-Jepsen, M., Kessing, L. V., & Bardram, J. E. (2013). Supporting disease insight through data analysis: Refinements of the monarca self-assessment system.

Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 133–142. https://doi.org/10.1145/2493432.2493507

Grist, R., Porter, J., & Stallard, P. (2017). Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. *Journal of Medical Internet Research*, *19*(5), e176. https://doi.org/10.2196/jmir.7332

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691616650285

Luxton, D. D. (2016). Chapter 1—An Introduction to Artificial Intelligence in Behavioral and Mental Health Care. In D. D. Luxton (Ed.), *Artificial Intelligence in Behavioral and Mental Health Care* (pp. 1–26). Academic Press. https://doi.org/10.1016/B978-0-12-420248-1.00001-5

Light, D. (2020, June 27). *New Prescription Drugs: A Major Health Risk With Few Offsetting Advantages*. Harvard Center for Ethics. https://ethics.harvard.edu/blog/new-prescription-drugs-major-health-risk-few-offsetting-advantages

National Institute of Mental Health *NIMH » Major Depression*. (n.d.). Retrieved November 10, 2020, from https://www.nimh.nih.gov/health/statistics/major-depression.shtml

Wisdom, J. P., Clarke, G. N., & Green, C. A. (2006). What Teens Want: Barriers to Seeking Care for Depression. *Administration and Policy in Mental Health*, *33*(2), 133–145. https://doi.org/10.1007/s10488-006-0036-4