EmoCure: Utilising Social Media Data and Smartphones to Predict and Cure Depression

Ashish Basantani, Yash Kesarwani, Srishti Bhatia, and Sarika Jain
National Institute of Technology, Kurukshetra

Abstract. Depression is a common and recurring illness. It is a serious medical condition which might lead to self-harm. Various factors like unawareness about depression, shortage of medical health professionals and social stigmas make proper treatment for depression inaccessible. With the growing importance of technology in every sector of life, healthcare systems have also started using technology to provide better treatment. Various studies have shown that the widespread use of smartphones can be useful in predicting as well as treating depression by recommending activities. Smartphones can be very useful in the continuous monitoring of a patient which in turn helps to keep track of the activities of the patient. Social media data can also be used to find out the mental state of a patient. We propose EmoCure, a smartphone application that uses social media data, wearable sensor data, patient history, and smartphone usage patterns to predict, monitor, and treat depression using emotion regulating activities. We use machine learning models for finding out the sentiments in the social media posts. To predict depression, we use ensemble learning. We then recommend personalized emotion regulating activities whichever the user prefers.

Keywords: Activity Recommendation, Depression, Mobile Sensors, Prediction, Social Media Data

1. Introduction

Depression has become a worldwide problem with more than 300 million people suffering from it, as reported in a research done by World Health Organization (WHO) 0. According to a study done by WHO, 7.5 percent of the Indian population suffers from some or other form of mental disorder. In the year 2015, India recorded the highest number of cases of depressive disorders in the world 0. Severe depression for a continuous period can also lead to suicide. Various healthcare workers are working consistently to find out a proper treatment of depression. There are various social stigmas attached with depression that inhibit the proper treatment of depression. The authors in 0 prove that social stigma is the prominent reason that leads to depression being untreated. Unawareness about disease is
also one of the factors. The rise in the number of people suffering from depression from 1990 to 2017 can be seen from the below image (Fig 1.1).

Earlier, the healthcare workers asked a patient to be physically present at their clinic for receiving any type of recommendations about health. Mostly, the suggestions that doctor gave were based on the description given by the patient, symptoms that he can observe and the lab test results. The symptoms however can be misleading or not very well described, due to which the physicians may infer wrongly and may not be able to properly predict the patient’s condition. The situation is even worse, where a patient has a mental disorder. There were no systems available earlier to keep track of the patient if he is not physically present at the doctor’s clinic. Medical experts will be able to provide better treatment to the patients, if the patient can somehow be monitored remotely and continuously.

The rapid advancement in mobile phone technology and use of social media has presented an opportunity for providing easily accessible and affordable healthcare. Mobile phones are devices that monitor our activities 24x7. Mental health behavior can be assessed by continuous monitoring of the patient via sensors and via social media. Various studies have been done that explore the idea of using mobile phones and social media for providing patient-centric healthcare.

The authors in 0 present a framework to use mobile phones for delivering healthcare. The use of smartphones has also given rise to the use of social media. There are around 300 million users of Facebook 0, 33 million users of Twitter and 26 million users of LinkedIn 0. People spend a lot of their time on social media daily. On an average, people spend 2.5 hours a day on social media globally. Social media platform has become a medium to share activities, feelings, etc. A lot of people now believe in sharing feelings, making friends, sharing common interest and ideas on virtual medium more than the real life. Facebook, twitter, reddit, etc. has become an outlet for the emotions for people. Thus, analyzing social media posts can be very helpful in determining the mental state of a person.
Munmun De Choudhury et al. used this approach for understanding depression in individuals and populations and found social media to be advantageous in curing depression.

We have studied the already existing systems which aimed to predict and cure depression. We found out the various gaps that were present in the existing systems. We then propose our system that performs prediction of depression as well as monitoring and activity recommendation for helping patients suffering from depression. We discuss various machine learning algorithms that we used for prediction of depression and which resulted into the best one. We also discuss the recommendation methodology that we are implementing into our system.

The proposed system aims to provide a cost-effective cure for depressive disorders using smartphones and social media data by continuous monitoring and recommending emotion regulating activities. The objectives of this study are as follows:

- To predict depression using smartphone usage patterns, patient history, wearable sensor data, and social media data.
- To monitor the patients in daily life by collecting their smartphone usage patterns, wearable sensor data, and social media data.
- To provide patient-centric healthcare and cure depression through self-therapy by using data collected during monitoring to recommend emotion regulating activities.

2. Related Work

With the advent of technology, mobile phones have become an important part of our day to day life. From listening music, surfing Internet to navigating to a new location, everything is possible using a smartphone. Smartphones giving so much options makes us carry it wherever we go. These devices can hence become a medium to provide patient-centric healthcare. The authors in introduce the concept of Augmented Personalized Healthcare (APH). According to APH, the prediction and treatment of depression requires continuous monitoring of a person and for this purpose a smartphone is the best option available.

A reason behind the extensive use of mobile phones is social media. Nowadays, people spend a lot of their time on social media websites such as Facebook, Twitter, Reddit and Instagram. People tend to share their thoughts and feelings on these virtual platforms more often than in real life. Analysis of a person’s social media data can be quite effective input in predicting their mental state. By doing continuous monitoring of a person’s activities, we can determine the existence of the key symptoms of depression, like isolation, decreased activity, sleeping problems, etc. Hence to predict the onset of depression in a person, smartphone usage patterns and smartphone sensor data can be effectively used. Apart from prediction, continuous monitoring through mobile phones can prove to be very useful in recommending activities too.

MoodScope developed by Microsoft, is a smartphone software to predict the mood of the user based on various inputs like text messages, emails, call logs, application usage, web browsing, and location. MoodScope is a mood logger application that is used by participants to self-report emotions. The data that is taken as an input is used to train a personalized mood model for each user using Multi-Linear
regression. The personalized model achieved an accuracy of 93%, while an All-User mood model that can be used as a general model for all users initially achieved an accuracy of 66%. The study also develops a hybrid model, which achieves an accuracy of 72%. Galen Chin-Lun Hung et al. 0 build their system using a similar approach. They used mobile phone usage patterns to train four classifiers which helped to predict negative emotions. The authors found that the Naïve Bayes classifier with greedy best forward feature selection had the highest accuracy of 86.71%.

A research done by Munmun De Choudhary et al. 0 uses crowdsourcing to collect tweets of individuals from Twitter who reported that they have been diagnosed with clinical depression. The authors trained a Support Vector Machine (SVM) classifier using various attributes such as the number of tweets per day, the number of replies, the pattern of tweets, linguistic styles, etc. The SVM classifier achieved an accuracy of ~70% with 0.74 precision. Working on similar approach Johannes C. Eichstaedt et al. 0 used Facebook posts to predict depression in medical records. They trained logistic regression models to achieve a high accuracy of 72%. The study 0 aims to develop a web application that provides the user an easy-to-use medium to check their depression levels using their social media posts. The web application also provides a doctor's location nearest to user’s location, if the person is found to be suffering from depression. Facebook posts along with questionnaires filled by the user was given as an input to Naïve Bayes classifier to detect depression. The authors 0 predict depression using social media data. They captured publicly available Facebook data of bipolar, depression, and anxiety pages that contained users’ comments using the NCapture tool. They train various classifiers and found that the Decision Tree classifier gives the best accuracy for their dataset. Gaur M et al. 0 used information (subreddits) from Reddit to check about suicidal tendencies and other mental health issues afflicting depressed users. The authors have built suicide risk severity lexicon using medical knowledge bases and suicide ontology to detect suicidal thoughts and actions. They found that for a 5-label confusion matrix CNN correctly classifies users with an accuracy of 80% and for a 4-label confusion matrix, it has 92% accuracy.

Galen Chin-Lun Hung et al. 0 developed an activity recommender system based on smartphone usage patterns, personal information, and environmental sensor information. The authors used clustering and combinational cube matrix (CCC) to recommend similar activities to similar users.

Intellicare is a suite of apps that aims to treat depression using interactive skill-based training that is based on cognitive-behavioral therapy, problem-solving therapy, etc. These apps were available on the Google Play Store. The study 0 evaluates the efficiency of Intellicare in reducing depressive symptoms. Shiqi Yang et al. 0 develop emHealth, an intelligent activity recommender system with depression prediction for emotion regulation. The authors first predict depression using decision tree technique and SVM as the prediction models and compare their accuracies. After detecting depression, the system recommends activities based on external factors of depression and the level of depression.

Fabian Wahle et al. 0 developed Mobile Sensing and Support (MOSS), a smartphone app that provides just-in-time interventions derived from cognitive behavior therapy. They use features such as the geographic movement, number of calls, time at home, number of calendar events, etc. to determine the actual behavior of the user and recommend activities accordingly.

The existing systems had a number of gaps. Firstly, the existing system use insufficient inputs to predict and cure depression. Most systems only use social media or smartphone usage patterns for
prediction and activity recommendation. Using only these inputs leads to inaccurate results. Secondly, most systems use Twitter for monitoring the social activity of the user. Using Twitter data may lead to fetching irrelevant posts on the basis of certain keywords. Most systems that we studied didn’t provide a single framework for both depression prediction and activity recommendations.

3. Methods

3.1. System Overview

The proposed system covers all the gaps in the existing systems. The system uses both social media data and smartphone usage patterns to predict and cure depression. The system also uses patient history and wearable sensor data. The system uses Reddit data instead of Twitter data to monitor social activity of the patients. Reddit is one of the most popular websites where people discuss topics, and there is a specific subreddit for each subject. It has a specific depression subreddit where most of the posts are from patients suffering from depression or mental health professionals. The system also provides a one-stop solution for predicting depression and activity recommendation.

The system comprises of three modules: Prediction, Monitoring and Recommendation. Figure 3.1 shows the system architecture. The Prediction module uses social media data, smartphone usage patterns, patient history and wearable sensor data predict whether the user is depressed or not. The Monitoring module is responsible for continuously monitoring the users’ current mood. The Recommendation module is responsible for recommending mood regulating activities to the user. When the user logs in to the EmoCure application, a dashboard containing three options i.e., Prediction, Monitoring, and Recommendation are shown. Figure 3 shows the dashboard of the EmoCure application. The user can click on these options to perform the corresponding activity.

![System architecture](Figure 3.1. System architecture)
The system performs prediction and activity recommendation using the following inputs:

- **Social Media Data**: The data that a person shares on their social media account can be used to figure out his state of mind. This data can be used effectively when the patient has a mental illness such as depression.
- **Patient History**: Patient’s history will be beneficial in recommending better activities to the patients. For example, a person who was diagnosed with depression in the past has higher chances of being depressed again.
- **Smartphone usage patterns**: Smartphone usage patterns such as call logs, text messages can describe how much the user interacts with people. They provide insights into the daily activity of the users.
- **Wearable Sensor Data**: Wearable sensors such as Fitbit describe the physical activity of the users, i.e., the number of steps they have walked, how much exercise they have done, heart rate, etc. Smartphones are also equipped sensors such as GPS sensors that describe the places they have visited in a day.

### 3.2. Data Collection & Pre-processing

The inputs provided to the system are collected in the following ways:

#### 3.2.1 Social media data

For training the depression prediction model, we collected the social media data from Reddit. Reddit is one of the most popular websites where people discuss topics, and there is a specific subreddit for each subject. The popularity and the content length make it ideal for prediction. It has a specific depression subreddit where most of the posts are from patients suffering from depression or mental health professionals. We use the Reddit Pushshift API to extract Reddit submissions from the depression subreddit. The API is available on the Internet free of cost.
To pre-process the Reddit posts, we need to find the overall sentiment of each post. The sentiment of each post is calculated by first tokenizing the post into words. The sentiment of the post is the summation of the polarity of the tokenized words divided by the total number of words in the post. To find the polarity of the words, we downloaded a dictionary which contains words of English language with their corresponding polarity i.e. negative or positive. For example, words like stress, sad and death are negative words whereas words like happy, excited and good are positive words. The overall sentiment of each post is calculated and stored with a unique ID assigned to the post in a CSV file. The dataset we collected contains 9331 Reddit posts.

3.2.2 Patient history data
For training the depression prediction model, we collected the patient history data from the Internet. The dataset contains attributes such as expenses, income, diseases, no of children, and other factors that can lead to depression for the farmers of South Africa.

The dataset initially had 76 attributes. We performed feature elimination manually by discarding attributes such as wage and livestock, which we found to be irrelevant. The dataset finally has 20 features. The dataset initially had 1143 instances. We discard those instances for which 50% of the attributes had missing or no values. The dataset finally has 1100 instances. The survey date is in the string format, so it is converted to a number using Label encoder.

The users’ patient history data is collected using the EmoCure mobile phone application. The user enters attributes such as age, working hours, income, medical conditions etc. which represent the history and the lifestyle of the patient.

3.2.3 Smartphone usage patterns
Smartphone usage patterns such as app usage, location data, calls, etc. are required for activity recommendation based on current context. The EmoCure android application collects this data while monitoring the user. Permissions for collecting and storing each type of data are taken from the user.

3.3. Prediction
For prediction that a person is suffering from depression or not requires the complete knowledge of the person, from his social background in the real world to his social presence in the virtual world. Various research have shown that the social media can be very useful tool in determining whether a person is suffering from depression or not. To track his real-world activities continuous monitoring of people can also be done using the smartphones. The key symptoms of depression, such as isolation, decreased activity, and sleeping problems can be detected by continuous monitoring of a person’s activities through wearable sensors and mobile phones.

We use ensemble learning to predict depression. Using ensemble learning, we can train different classifiers for the inputs and combine their outputs to give the final prediction. This enables us to use classifiers that are best suited for individual inputs. Figure 3.3 represents the depression prediction module.
3.3.1 Social Media Data
We use machine learning methods to train the depression prediction model using social media data. We fetched the Reddit posts and their corresponding sentiment using the unique ID of each post for training the depression prediction model using social media data. We trained models using five different machine learning techniques, namely Decision Tree, Random Forest, Naïve Bayes, k-Nearest Neighbour, and Support Vector Machine. We found that of all the five algorithms, the decision tree gives the best accuracy. The Naïve Bayes method gave an accuracy of 93.79%, Support Vector Machine gave 50.0%, K-Nearest Neighbour gave 81.46%, Random Forest gave 49.1%, and Decision Tree gave the best accuracy of 98.55%. The sentiment analysis process is shown in Figure 4.4.

3.3.2 Patient History Data
For training the classifier for patient history data, we use neural networks. We use neural networks because traditional machine learning algorithms might fail in the future when the dataset grows in size, and neural networks can efficiently handle large data. The neural network for prediction has two hidden layers, with thirty-six neurons in each layer. We use 70% of the rows in the dataset for training.
and 30% for testing. The train_test_split python function does this for us. With 100 epochs, the model achieved an accuracy of 86.7%.

3.4. Monitoring

Monitoring the users’ mood is essential for tracking and controlling symptoms of depression in the user. It helps to understand the actual situation of the user. The MoodCheck module in the EmoCure application is responsible for monitoring the user. The MoodCheck module sends a notification to the user every four hours and asks her to input her current mood. The user can select a mood rating from 1 to 10 as shown in Figure 4.5. Mood information for every user is stored in the database. The MoodCheck module also shows the mood graph of the user for the day. The mood graph helps the users to track their own progress and their moods over the day. The user can also view their mood graph of any previous dates.

Figure 3.5. Enter user mood rating

Figure 3.5 shows the mood graph. Based on the current context and the current mood of the user, the application recommends mood regulating activities to the user.

Figure 3.6. Mood graph
3.5. Activity Recommendation

The system performs activity recommendations based on the user's interaction with the EmoCure application. The activity recommendation module takes into account the current context, user’s current mood and user’s past preferences. Initially, the activities are recommended randomly from the set of activities that psychiatrists recommend to their patients suffering from depression. After days of monitoring and collecting user input, the system recommends the activities which were preferred by the user in the past. Figure 4.7 represents the activity recommendation module.

Figure 3.7. Recommendation module

Whenever the user gives a negative mood input to the MoodCheck module, the app recommends mood regulating activities based on the current context and user’s past preferences. When the user gives a positive mood input, the application shows positive messages and recommends light activities such as listening to happy music.

For storing user preference data, a table is maintained for the activities in the database. The rows of the table depict the activities that can be recommended, and the columns depict one-hour time slots in a day. Each entry in the table corresponds to the number of times the user has chosen the particular activity during that particular time slot. Every time the user selects an activity the corresponding entry in the table increases. At the time of recommending an activity, the application will first fetch activities based on the maximum value in the table for the current time. The application will then perform contextualization by checking for multiple factors that affect the recommendation. Factors like environmental factors such as current weather, personal factors such as time constraints and working hours etc. represent the current situation of the user. These factors determine what activity can be recommended. For example, outdoor activities cannot be recommended when the current weather is rainy or snowy. Table 1 represents some of the activities and the factors that determine whether that activity can be recommended or not.
Table 1. List of activities

| S.No. | Activity                      | Description                                                                 | Conditions                                                                 |
|-------|-------------------------------|-----------------------------------------------------------------------------|----------------------------------------------------------------------------|
| 1.    | Social support                | Suggest to connect with favorite contacts.                                  | • Calendar is free for the current date and time.                          |
| 2.    | Self-affirmations             | Suggest practicing self-affirmations, read positive quotes/articles or life lessons. |                                                                            |
| 3.    | Exercise for mood             | Suggest exercises.                                                          | • Non-working hours.                                                      |
|       |                               |                                                                             | • Calendar is free for the current date and time.                          |
|       |                               |                                                                             | • Weather conditions are favorable.                                        |
|       |                               |                                                                             | • No disabilities/conditions that forbid exercise.                         |
| 4.    | Behavioral activation         | Track positive streaks, mood history.                                       |                                                                            |
| 5.    | Mindfulness practice          | Suggest different types of meditations.                                     | • Non-working hours.                                                      |
|       |                               |                                                                             | • Calendar is free for the current date and time                           |

After fetching highly preferred activities and performing contextualization based on user’s current situation and lifestyle, the system will recommend three activities to the user. The user can choose any which he/she finds comfortable to do.

Figure 3.8 shows a screenshot of the application. The user can tap on any of the recommended activities and then page for that activity opens up and the user can start performing the activity. The cycle of recommending the activity and choosing is continued and performed every four hours.
4. Results and Findings

The benchmarks systems that we have considered to predict depression use twitter data. The benchmark system 0 used a Support Vector Machine classifier and achieved 70% accuracy in depression prediction.

![Figure 4.1. Prediction results](image)

For depression prediction using patient history, we trained an Artificial Neural Network that gave an accuracy of 86.7% in 100 epochs. For depression prediction using social media data, we trained our model on five different classifiers. The Naïve Bayes classifier gave 93.79% accuracy, Support Vector Machine classifier gave an accuracy of 50.0%, k-Nearest Neighbour gave 81.46% accuracy, Random Forest gave an accuracy of 49.1% and Decision Tree classifier came out to be the best among all with an accuracy of 98.55%. Various results obtained by training our social media models are shown in Figure 4.1.

The confusion matrices for the different models trained for social media data is shown in Figure 4.2. The confusion matrix is a tabular representation of performance of a machine learning classification algorithm. These confusion matrices can be used to calculate the accuracy, precision and recall of a classifier. We can use the Sklearn library of the python to calculate these values. The columns represent the predicted label while the rows represent the true label for the posts. 1 represents positive post, 0 represents a neutral post while -1 represents a negative post. Each cell [i, j] (where i is the row and j is the column) in the matrix represents the number of posts with the true label represented by row i and the predicted label is represented using row j.
Figure 4.2. Confusion matrices

The confusion matrix for NB classifier tells that it classifies 2000 posts as positive sentiment posts correctly, while it wrongly predicted 118 positive posts as neutral posts and 76 posts as negative. The KNN classifier confusion matrix shows that it wrongly predicted 583 positive posts as negative and 185 positive posts as neutral posts. It predicted 1496 post correctly as having positive emotions. The confusion matrix of SVM classifier explains that it wrongly predicted 2264 positive posts as negative posts, and 0 posts as neutral & positive posts. The Random forest classifier correctly predicted 0 positive polarity posts as positive while 2146 positive posts as having negative polarity and 118 others as having neutral polarity. The Decision Tree classifier came out to be the best. Its confusion matrix says that it wrongly predicted 25 positive polarity posts as having negative emotion and 44 posts as neutral posts. It correctly predicted 2195 positive posts as having positive polarity better than all other classifiers.

5. Conclusion
Depression, which has become a global problem requires a cost-effective and accurate solution. We developed EmoCure mobile phone application that uses the idea of self-therapy to cure depression. The application is a step towards personalized patient-centric healthcare that is easily accessible and affordable. We use a unique combination of inputs that hasn’t been used in previous research for prediction and recommendations. We successfully predicted depression using machine learning methods and performed personalised activity recommendations using smartphone usage patterns, wearable sensor data, social media data, and Electronic Health Records of the users. The prediction accuracy achieved and the effectiveness of the system in curing depression prove the efficacy of digital healthcare. Therefore, mental healthcare apps like EmoCure have the potential to treat and manage depression through activity recommendations.
6. Future Scope
The EmoCure app is only available for Android mobile phones currently. We would like to develop the same for the iOS platform too. In the future, we would like to improve the social media data analysis by using aspect-based sentiment analysis techniques. Enhancing social media data analysis will lead to more accurate predictions. We would also like to improve the activity recommended by recommending activities in accordance with the level of depression the patient is suffering from.

References
[1] WHO Publications, https://www.who.int/publications-detail/depression-global-health-estimates, last accessed 2020/4/13. Another reference
[2] Chaturvedi, S K, Rao, G P, Sarda, K D, & Suryawanshi, S Y 2014 Chronic pain and depression: an online survey on Indian experiences Indian Journal of Pain 28(3) 166.
[3] Aung, M H, Matthews, M, Choudhury, T 2017 Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies Depression and anxiety 34(7) 603-609.
[4] Number of Internet users in India from 2015 to 2018 with a forecast until 2023, https://www.statista.com/statistics/255146/number-of-Internet-users-in-india/, last accessed 2020/4/13.
[5] Number of facebook users in India from 2015 to 2018 with a forecast until 2023, https://www.statista.com/statistics/304827/number-of-facebook-users-in-india/, last accessed 2020/4/13.
[6] De Choudhury, M, Gamon, M, Counts, S, Horvitz, E 2013 Predicting depression via social media Seventh international AAAI conference on weblogs and social media.
[7] De Choudhury, M, Counts, S, & Horvitz, E 2013 Social media as a measurement tool of depression in populations Proceedings of the 5th Annual ACM Web Science Conference pp 47-56.
[8] Sheath, A, Jaimini, U, Thirunarayan, K, Banerjee, T 2017 Augmented personalized health: How smart data with iots and ai is about to change healthcare IEEE 3rd International Forum on Research and Technologies for Society and Industry (RTSI) pp 1-6.
[9] LiKamWa, R, Liu, Y, Lane, N D, Zhong, L 2013 Moodscope: Building a mood sensor from smartphone usage patterns Proceeding of the 11th annual international conference on Mobile systems, applications, and services pp 389-402 ACM.
[10] Hung, G C L, Yang, P C, Chang, C C, Chiang, J H, Chen, Y Y 2016 Predicting negative emotions based on mobile phone usage patterns: an exploratory study JMIR research protocols 5(3) e160.
[11] Eichstaedt, J C, Smith, R J, Merchant, R M, Ungar, L H, Crutchley, P, Preo?iuc-Pietro, D, Schwartz, H A 2018 Facebook language predicts depression in medical records Proceedings of the National Academy of Sciences 115(44) 11203-11208.
[12] Sonawane N., Padmane M, Suralkar V, Webel S, Date P 2018 Predicting depression levels using social media posts International Journal of Innovative Research in Science, Engineering and Technology 7 Issue 5.
[13] Islam, M R, Kabir, M A, Ahmed, A, Kamal, A R M, Wang, H, Ulhaq, A 2018 Depression detection from social network data using machine learning techniques Health information science and systems 6(1) 8.
[14] Gaur, M, Alambo, A, Sain, J P, Kursuncu, U, Thirunarayan, K, Kavuluru, R, Pathak, J 2019 Knowledge-aware assessment of severity of suicide risk for early intervention The World Wide Web Conference pp 514-525.
[15] Hung, G C L, Yang, P C, Wang, C Y, Chiang, J H 2015 A smartphone-based personalized activity recommender system for patients with depression Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare pp 253-257.
[16] Mohr, D C, Tomasino, K N, Lattie, E G, Palac, H L, Kwasny, M J, Weingardt, K, Caccamo, L 2017 IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety Journal of medical Internet research 19(1) e10.
[17] Yang, S, Zhou, P, Duan, K, Hossain, M S, Alhamid, M F 2018 EmHealth: towards emotion health through
depression prediction and intelligent health recommender system *Mobile Networks and Applications* 23(2) 216-226.

[18] Wahle, F, Kowatsch, T, Fleisch, E, Rufer, M, Weidt, S 2016 Mobile sensing and support for people with depression: a pilot trial in the wild *JMIR mHealth and uHealth* 4(3) e111.

[19] https://github.com/niquejoe/Classification-of-Depression-on-Social-Media-Using-Text-Mining, last accessed 2020/05/05.

[20] https://zindi.africa/competitions/busara-mental-health-prediction-challenge/data, last accessed 2020/05/05.

[21] Bhatia S, Kesarwani Y, Basantani A, Jain S 2020 Engaging Smartphones and Social Data for Curing Depressive Disorders: An Overview and Survey *Paradigms of Computing, Communication and Data Sciences (PCCDS)*.

[22] Dalal S, Jain S, Dave M 2019 A Systematic Review of Smart Mental Healthcare, *Proceedings of the 5th International Conference on Cyber Security and Privacy (ICCS)* pp 38-44.

[23] Mishra S, Jain S, Abraham A, Shandilya S Dec 2018 Secure Semantic Smart HealthCare (S3HC) *Journal of Web Engineering, River Publishers* 17:8 pp 617-646.