International Journal of Research in Advent Technology, Vol.6, No.10, October 2018
E-ISSN: 2321-9637
Available online at www.ijrat.org

Psychosocial Problem and Depression Prediction using Social Media

D.Sridhar[#1], Dr.V.Kathiresan[*2]
[#]*Assistant Professor, Department of Computer Science, Dr.SNS Rajalakshmi College of Arts and Science, Coimbatore*
[*2] *HoD, Department of Computer Applications(PG), Dr.SNS Rajalakshmi College of Arts and Science, Coimbatore*
[1] *ursridhardeva@gmail.com,* [2]*vkathirmca@gmail.com*

Abstract- Mental health has turned into a general worry of individuals these days. It is of indispensable significance to identify and oversee mental health issues before they transform into extreme issues. Customary mental mediations are dependable, yet costly and hysteretic. With the quick advancement of social media, individuals are progressively offering their day by day lives and associating to companions on the web. Through collecting social media information, completely consider the location of mental health, with two average mental issues, depression and stress, as particular models. Introducing with paired user level location, we develop our examination towards various settings, by thinking about the trigger and level of mental health issues, and including distinctive social media stages of various societies. We develop benchmark genuine datasets for examination and propose a progression of multi-modular location models, whose adequacy is checked by broad tests with the utilization of Novel advance Ada boost algorithm. We additionally make a top to bottom examination to uncover the fundamental online practices in regards to these mental health issues. The decision is still out on whether social media is harming to the mental health of youngsters. This is partly because of the absence of research. A few examinations demonstrate that online associations with little gatherings of individuals can be gainful to youngsters, while other research focuses on an ascent in indications of uneasiness, dejection, and dietary problems. The other reason it's hard to get a decent read on the issue is that social media is always showing signs of change and developing. Also, no long haul considers have been finished. Along these lines, we're left making taught surmises dependent on momentum look into. There's sufficiently not information to back up the potential long-haul upsides and downsides of living on "likes.

Index Terms- Depression, behaviour, Advanced AdaBoost, mental health, Facebook, words, posts, likes

1. INTRODUCTION

With the quick pace of life, mental health has gotten broad consideration these days. Mental, neurological, and substance utilize disarranges add to more than 10 percent of the worldwide weight of illness. Normal side effects like pressure or clinical issue like melancholy, over the top and interminable bothersome mental states are very hurtful, and therefore it is of imperative centrality to identify mental health issues before they prompt serious results. Numerous utilizations of social media include data mining, for example, understanding user interests, client surveys, and supposition around news occasions. We examine utilization of social media content mining with regards to understanding open mental health.

This is an increasingly important application domain:

The commonness of mental health conditions is expanding, as is the measure of information we need to comprehend these conditions. While content information has been dissected in incredible profundity for marketing purposes, there remains an expansive open door in utilizing content information to comprehend open mental health. A few analysts have distinguished social media information, for example, FB posts, as a profitable hotspot for mental health signals. In any case, there stay basic holes in our capacity to comprehend mental health. Individuals may not share mental health content openly, particularly on the biggest social media stages, which are related to an individual character. In this paper, we play out a mental health examination on defamed themes by exploiting an interesting dataset from a social media for posting diary entries and sharing and following states of mind. We find issues that are not broadly talked about on other social media, for example, rest.

Mental health conditions influence a substantial level of people every year. Conventional mental health contemplates have depended on data gathered through contact with the mental health professional. There has been a study on the utility of social media for discouragement, however, there have been restricted assessments of other mental health conditions (Jan-Are, Jan and Deede, 2002). In the first place, we will inspect particular procedures that have already been utilized to examine gathering information, social health, and general health issues, and in conclusion, we will investigate the suggestions that this examination has for huge information examination.

There are some positive perspectives on social media. Remember that youngsters are designed for socialization, and social media makes socializing simple and immediate. Teenagers who battle with social abilities, social tension, or who don't have simple access to eye to eye socializing with different adolescents may profit by interfacing with different youngsters through social media. Teenagers in underestimated bunches including LGBTQ (Lesbian Gay Bisexual Transgender Queer) adolescents and youngsters battling with mental health issues can discover support and companionship through utilization of social media. At the point when teenagers interface with little gatherings of steady youngsters by means of social media, those associations can be the distinction between living in seclusion and inventing support.

2. BACKGROUND STUDY

Social media are utilized as a methods by which youth utilize their requirement for socialization (Gosling et al., 2011; Wang et al., 2014). For instance, it has been discovered that Internet habit and dependence on social systems have a solid relationship, notwithstanding, extraversion is a trademark that can show up in continuous clients of social systems (Müller et al., 2016). Albeit independent individuals are relied upon to probably utilize social systems for their correspondence needs, some different factors, for example, the littler size of their social system in reality and low level of self-assurance might be a burden in them. (Wang et al., 2014). In spite of the fact that not really identified with enslavement, another deciding component identified with the utilization of social media is thoughtfulness. Individuals who have a tendency to be more lovely have a tendency to have more interest in their pages and in those of others than individuals with low consideration (Gosling et al., 2011).

As of late the examination on the utilization and enslavement of Facebook has centered (Błachnio et al., 2017; Hong et al., 2014), in any case, it is intriguing to separate between social systems (Kuss and Griffiths, 2017). For instance, contrasts have been found between the powerlessness of enslavement in Instagram clients, in contrast to FB clients (Donelly and Kuss, 2016). In any case, a few inquiries likewise emerge in regards to dependence on innovation. It is viewed as that the utilization of innovation is somewhat an implies that enables individuals to do specific practices, for example, the utilization of social systems. Consequently, the reliance on social systems is identified with its substance and not to the innovation itself (Kuss and Griffiths, 2017). This corresponds with the hazardous utilization of the cell phone, where it could be connected more as a method for access than versatile reliance itself (Jasso et al., 2017).

Cell phones are right now a standout amongst the most prominent methods for access, with 90% of individuals getting to by this implies. Youngsters are the primary individuals of these gadgets, turning into a vital part and an augmentation of their lives, expanding their correspondence and social connections, yet in addition promoting issues and unhealthy propensities for utilize, including addictive conduct (Kwon et al., 2013; Hong, Chiu and Huang, 2012; Roberts, Pullig and Manolis, 2015). The utilization of cell phones is viewed as a conceivable indicator of the addictive conduct of social media because of their availability to the web (Echeburúa and de Corral, 2010; García, 2013; Orsal, Orsal, Unsal and Ozalp, 2013; SeHoon, HyoungJee, Jung-Yong and Yoori, 2016). For instance, there have additionally been contrasts in the dependence on the cell phone in connection to content, with social systems being a more prominent indicator than the utilization of portable diversions (Se-Hoon et al., 2016). Since getting to social media has turned out to be a standout amongst the most well-known exercises completed on the web, there is expanding worry about dangerous or addictive examples with respect to the utilization of cell phones; it is considered as a specific developing issue in youngsters (Lam, 2014, 2015; Lee, 2015; Yau et al., 2014).

The Downside Social Media Use by Teens

Read enough of the ebb and flow research and you'll see that the negatives tend to feel greater than the positives. While youngsters can utilize social media to interface and make kinships with others, they additionally stand up to digital tormenting, trolls, dangerous correlations, lack of sleep, and less incessant eye to eye connections, to give some examples. An excessive amount of time spent looking through social media can result in indications of nervousness and additionally discouragement. Here's the manner by which social media can be ruinous:

- **Focusing on likes**: The need to pick up "likes" on social media can make teenagers settle on decisions they would some way or another not make, including adjusting their appearance, participating in negative practices, and tolerating unsafe social media challenges.
- **Cyberbullying**: Youngsters specifically are in danger of digital tormenting through utilization of social media, however, teenager young men are not resistant. Digital tormenting is related to despondency, nervousness, and a raised danger of self-destructive considerations.
- **Making comparisons**: Though In spite of the fact that numerous youngsters realize that their companions share just their feature reels on social media, it's exceptionally hard to

abstain from making examinations. Everything from physical appearance to life conditions to apparent triumphs and disappointments is under a magnifying instrument on social media.

- **Having too many fake friends**: Indeed, even with protection settings set up, teenagers can gather a huge number of companions through companions of companions on social media. The more individuals on the companion list, the more individuals approach screen capture photographs, Snaps, and updates and utilize them for different purposes. There is no protection on social media.
- **Less face time**: Social cooperation aptitudes require every day rehearse, notwithstanding for teenagers. It's hard to fabricate sympathy and empathy (our best weapons in the war on harassing) when adolescents invest more energy "drawing in" online than they do face to face. The human association is an intense apparatus and fabricates aptitudes that endure forever.

3. PROPOSED METHODOLOGY

We created computational models to anticipate the rise of gloom and Post-Traumatic Stress Disorder in FB users. FB information and points of interest of discouragement history were gathered from 50 users during August, September and October 2017. We separated prescient highlights estimating influence, semantic style, and setting from member and fabricated models utilizing these highlights with novel Advanced AdaBoost algorithm. Coming about models effectively segregated among discouraged and healthy substance, and contrasted positively with general specialists' normal achievement rates in diagnosing dejection, though in a different populace. Results held notwithstanding when the investigation was limited to content posted before first gloom conclusion. State-space transient investigation proposes that beginning of despondency might be perceivable from FB information a while preceding finding. Prescient outcomes were repeated with a different example of people determined to have wretchedness. These techniques propose an information-driven, the prescient methodology for early screening and identification of mental sickness.

Data Set

Members were selected utilizing FB stage, and gathered client information from FB history. Enlistment and information gathering techniques were indistinguishable for depression samples. Members were requested to share their FB usernames and history. After anchoring assent, we made a one-time gathering of members' FB posting history, up to the most recent post. In absolute we gathered 9,524 unique information from 50FB clients for the misery investigation. Points of interest on member information insurance measures are sketched out underneath

Attributes

- Age
- Gender
- Posting frequency
- Comments
- Like Pages
- Search History
- Check ins
- Login and Logout time

Participant safety and privacy

This investigation configuration raised two essential issues in regards to moral research hones, as it concerned the two people with mental disease and conceivably actually identifiable data. We were not able assurance strict obscurity to members, given that usernames and individual data presented on FB are frequently naturally particular to members' personalities, (for example, usernames and posts containing genuine names). As we conceivably had the ability to connect think about members' characters to delicate health data, contemplate users were educated of the dangers of being by and by recognized from their social media information. Users were guaranteed that no closer to home identifiers, including usernames, could ever be made open or distributed in any method.

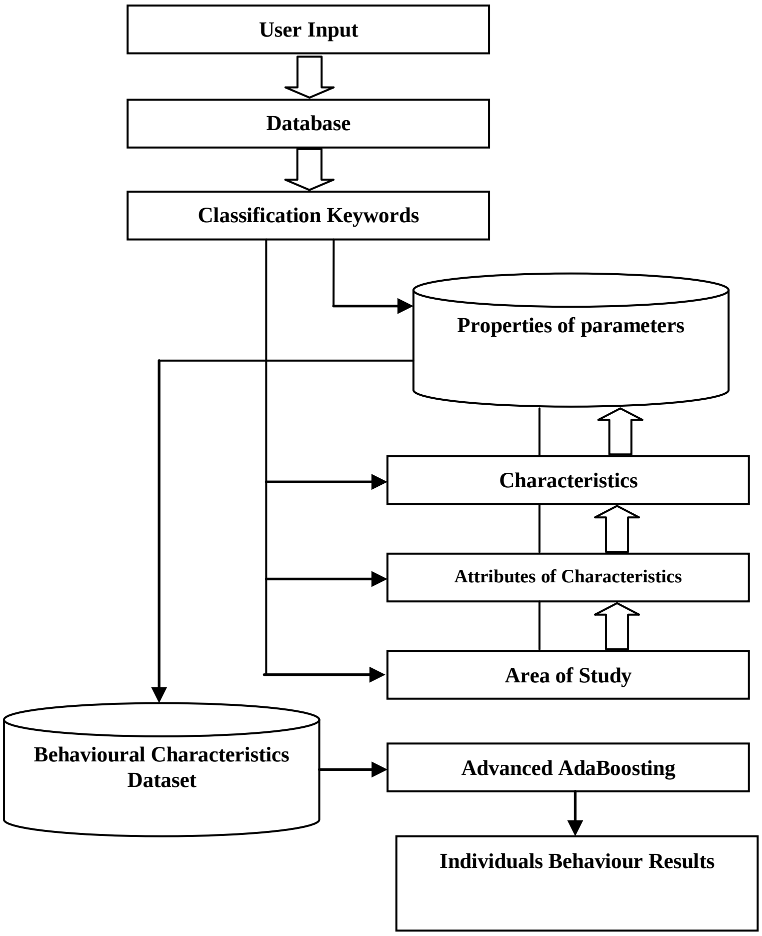

Figure 1: Architecture of Advanced AdaBoost model for predicting social media user's mental health

Algorithm
AdaBoost stands for adaptive boosting and it assumes that finding many weak models are easier than finding one accurate model. Here we are using Advanced AdaBoost algorithm since Ada boost algorithm is sensitive to noisy data and outliers. We have added some features like labelled data from humans, Gain insights from manual error analysis, etc. Boosting is an approach to create predictions rules with high accuracy using a combination of weak models and rules that have low accuracy in prediction. Boosting generates a sequence of base models and then decides a final estimate of the target variable based on aggregating the estimates of the base models. Advanced AdaBoost generates a numbers of weak classifiers and a final estimate of the target variable is chosen based on aggregating the estimates made by the base models. Similar to the random forest algorithm, Advanced AdaBoost also have a variable importance estimation but in a different way. In Advanced AdaBoost the more informative variables are used more often, and the less informative features are barely used.

Advanced AdaBoost (data, learner):

1. $N \leftarrow nrow(data)$
2. $M \leftarrow ncol(data)$
3. $x \leftarrow data[,1:M-1]$
4. $y \leftarrow data[,M]$
5. $for\ i \leftarrow to\ N{:}\ \omega_i = 1/N$
6. $Repeat\ i \leftarrow 1, i \leftarrow i+1{:}$
7. $M_i \leftarrow learner\ (data, \omega)$
8. $ms = \{p | M_i(x_p) \neq y_p\}$
9. $\in_i = \frac{\sum_{i \in ma} \omega_j}{\sum_{j=1}^{n} \omega_j}$
10. $a_i = \log((1 - \in_i)/\in_i)$
11. $for\ j \in ms{:}\ \omega_j = \omega_j * e^{ai}$
12. $for\ i \leftarrow 1\ to\ N{:} \omega_j = \frac{\omega_j}{\sum_{j=1}^{n} \omega_j}$
13. $until \in_i \geq 0.5\ or\ ms = \emptyset$
    $Return\ [M(x) =$
14. $sign\left(\sum_{j=1}^{T} a_j M_j(x)\right)]$

The final model M (line 14) combines the other models using a weighted sum of the outputs of these other models. The weights a_i reflect the accuracy of each of the constituent models.Now we can illustrate the process. The number of training entities N is 50. Each weight ω_j is thus initially 0.1 (line 5). Imagine the first model, M_1, correctly classifies the first 30 of the 50 entities (sample:$ms = \{7,8,9,10\}$), so that $\in_1 = 0.1 + 0.1 + 0.1 + 0.1/1 = 0.4$. Then$a_1 = \log\left(\frac{0.6}{0.4}\right) = 0.405$, and is the weight that will be used to multiply the results from this model to be added into the overall model score. The weights$w_7, \ldots, w_{10}$ then become, $0.1 * e^{0.405} = 0.1 * 1.5 = 0.15$. That is, they now have more importance for the next model build. Suppose now that M_2, correctly classifies 8 of the entities (with ms={1, 8}), so that $\in_2 = \left(01 + \frac{0.15}{1.2}\right), \ldots 0.208\ and$ $a_2 = \log(0.792/0.208)$.

Thus$w_1 = 0.18 * e^{1.337} \approx 0.381$ and $w_8 = 0.15 * e^{1.337} \approx 0.571$. Note how record 8 is proving particularly troublesome and so its weight is now the highest.We can understand the behaviour of the function used to weight the models$\left(\log\left(\frac{(1-\in_i)}{\in_i}\right)\right)$ and to adjust the observation weights $e^{ai}$by plotting both of these for the range of errors we expect (from 0 to 0.5 or 50%).

4. RESULTS FOR PREDICTING SOCIAL MEDIA USER'S MENTAL HEALTH

All data collection took place during August, September and October 2017. Across both depressed and healthy groups, we collected data from 50 FB users, totalling9,524posts. This number includes up to 200 posts from each participant's FB history; the analyses in this report focus only on posts from depressed users created before the date of first depression diagnosis. The mean number of posts per user was 1372 (SD = 1281.74). This distribution was skewed by a high number of frequent posters, as evidenced by a median value of just 375 posts per user. Table 1 discussed summary statistics.

Table 1. Summary statistics for depression post collection (Ndepr = 9524)

Depression	Users	Post	Posts (μ)	Posts (median)
Total	50	9524	1372	287
Depressed	32	5700	621	375
Healthy	15	3785	389	191
others	3	39	362	107

S.No	Word	+/-'	Up/Down	%Cont
1	Picture	+	Down	-5.46
2	Can't	-	Up	-3.59
3	no	-	Up	-3.18
4	not	-	Up	-2.61
5	kill	-	Up	-2.25
6	like	+	Down	1.55
7	park	+	Down	2.69
8	spoiled	-	Up	3.58
9	death	-	Up	-4.32
10	jail	-	Up	-1.98
11	I	+	Down	2.25
12	never	-	Up	-3.14
13	happy	+	Down	3.9
14	sad	-	Up	-0.5
15	shit	-	Up	1.9

5. *Feature extraction*

Depressed FB users have been observed to posts frequently than non-depressed users, so we used total posts per user, per day, as a measure of user activity. Posts metadata was analyzed to assess average word count per post (here, a word is defined as a set of characters surrounded by whitespace), whether or not the posts was a reposts, and whether or not the posts was a reply to someone else's posts. The labMT and LIWC 2007 were used to quantify the happiness of posts language. The use of labMT, which has shown strong prior performance in analyzing happiness on FB, is novel with respect to depression screening; LIWC have been successfully applied in previous studies on depression and FB. LIWC was also used to compile frequency counts of various parts of speech (e.g., pronouns, verbs, adjectives) and semantic categories (e.g., food words, familial terms, profanity) as additional predictors.

Table 2: Top Depression Predictors

Mood	Word Count
Happy	18
arousal	10
dominance	11
ingest	8
sad	14
swear	8
article	4
posemo	3

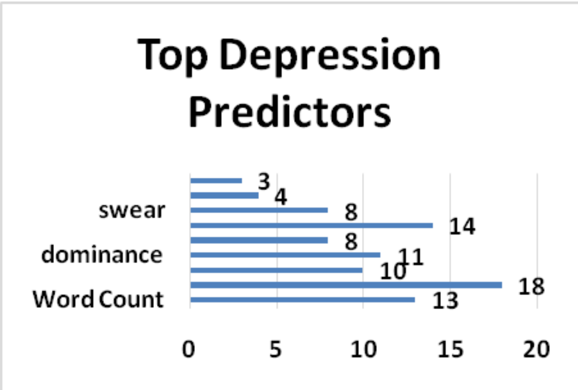

Figure 2: Top Depression Predictors

For depression samples ($N_{depr} = 9524$). Predictor names ending in "happy" are happiness measures; LIWC predictors refer to the occurrence of semantic categories.

Table 3: Word shift graphs

We averaged labMT happiness scores across observations in each class, after the removal of common neutral words and re-posted promotional material. Neutral words were words with labMT happiness scores between 4 and 6, on a 1–9 scale.

This included many common parts of speech, including articles and pronouns, which contributed little to understanding inter-group differences in valenced language. Some of the positive language observed more frequently among healthy individuals came from re-posts of promotional or other advertising material (like "win", "free", "gift"). We removed obvious promotional reposts when generating word shift graphs, as their removal did not significantly change mean posts-happiness differences between groups, and the resulting graphs gave better impressions of what participants personally posted about. We observed that postss authored by the depressed class were sadder ($h_{avg} = 6.01$) than the healthy class ($h_{avg} = 6.15$). In Fig. 4, we rank order individual words with respect to their contribution to this observed difference, and display the top contributing words

Figure 3: Word Shift Graph

Depression word-shift graph revealing contributions to difference in FB happiness observed between depressed (5.98) and healthy (6.11) participants. In column 3, (−) indicates a relatively negative word, and (+) indicates a relatively positive word, both with respect to the average happiness of all healthy posts. An up (down) arrow indicates that word was used more (less) by the depressed class. Words on the left (right) contribute to a decrease (increase) in happiness in the depressed class.

Precision : Positive prescient esteem is the part of pertinent occurrences among the recovered cases. Accuracy is the quantity of the right element separated by the quantity of all returned highlight space.

$$\text{Precision} = \frac{\text{True positive}}{Truepositive + FalsePositive}$$

Recall : It is the portion of pertinent cases that have been recovered over the aggregate sum of important cases. The review is the piece of the applicable reports that are effectively arranged into the correct classes

$$\text{Recall} = \frac{\text{True positive}}{Truepositive + Falsenegative}$$

F-Measure: It is the number of correct class predictions to the single document to total number of predictions to whole document.

Accuracy is given by

$$\frac{\text{True positive} + \text{True Negative}}{Truepositive + TrueNegative + falsepositive + F}$$

Table 4: Performance of advanced adaboost algorithm

Technique	Precision	Recall	F measure
Advanced AdaBoost**method for social media mental health prediction**	0.85	0.91	0.81

5. CONCLUSION

As the methods employed in the present study aim to infer health related information about individuals, some additional cautionary considerations are in order. Data privacy and ethical research practices are of particular concern, given recent admissions that individuals' Facebook and dating profile data were experimentally manipulated or exposed without permission. Indeed, we observed a response rate reflecting a seemingly reluctant population. Future research should prioritize establishing confidence among experimental participants that their data will remain secure and private. Complicating efforts to build socio-technical tools such as the models presented in this study, data trends often change over time, degrading model performance without frequent calibration55. As such, our results should be considered a methodological proof-of-concept upon which to build and refine subsequent models.This report provides an outline for an accessible, accurate, and inexpensive means of improving depression, especially in contexts where in-person assessments are difficult or costly. In concert with robust data privacy and ethical analytics practices, future models based on our work may serve to augment traditional mental health care procedures. More generally, our results support the idea that computational analysis of social media can be used to identify major changes in individual psychology.

REFERENCE

[1] Chancellor S, Lin Z, Goodman E, Zerwas S, De Choudhury CM. Quantifying and predicting mental illness severity in online pro-eating disorder communities. 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; Feb 27-Mar 2, 2016; San Francisco, CA, US. 2016. pp. 1171–1184.

[2] De Choudhury M, Counts S, Horvitz E. Social media as a measurement tool of depression in populations. 5th Annual ACM Web Science Conference; May 2-4, 2013; Paris, France. 2013. pp. 47–56.

[3] Gupta R, Malandrakis N, Xiao B, Guha T, Van SM, Black M, Potamianos A, Narayanan S. Multimodal prediction of affective dimensionsand depression in human-computer interactions. 4th International Workshop on Audio/Visual Emotion Challenge - AVEC ’14; Nov 7, 2014; Orlando, FL, USA. 2014. pp. 33–40.

[4] Hao B, Li L, Gao R, Li A, Zhu T. Sensing subjective well-being from social media. In: Slezak D, Schaefer G, Vuong ST, Kim YS, editors. Active Media Technology. AMT 2014. Lecture Notes in Computer Science. Volume 8610. Cham, Switzerland: Springer; 2014. pp. 324–335.

[5] Homan C, Johar R, Liu T, Lytle M, Silenzio V, Ovesdotter AC. Toward macro-insights for suicide prevention: analyzing fine-grained distress at scale. Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Jun 27, 2014; Baltimore, MD, USA. 2014. pp. 107–117.

[6] Huang X, Zhang L, Chiu D, Liu T, Li X, Zhu T. Detecting suicidal ideation in Chinese microblogs with psychological lexicons. 2014 IEEE International Conference on Scalable Computing and Communications and its Associated Workshops; Dec 9-12, 2014; Bali, Indonesia. 2014.

[7] Khadjeh Nassirtoussi A, Aghabozorgi S, Ying Wah T, Ngo DCL. Text mining for market prediction: a systematic review. Expert Syst Appl. 2014 Nov;41(16):7653–7670. doi: 10.1016/j.eswa.2014.06.009.

[8] Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D. Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines. Am Psychol. 2015 Sep;70(6):543–56. doi: 10.1037/a0039210.

[9] Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D. Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines. Am Psychol. 2015 Sep;70(6):543–56. doi: 10.1037/a0039210.

[10] Preotiuc-Pietro D, Sap M, Schwartz H, Ungar L. Mental illness detection at the World Well-Being Project for the CLPsych 2015 shared task. 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 5, 2015; Denver, CO, USA. 2015. pp. 40–45.

[11] D.Sridhar, Dr.V.Kathiresan. Predicting Mental Behaviour of Social Media Users Using AOMM (Auto Observation Mining Method), Jour of Adv Research in Dynamical & Control Systems , Jan 2018

[12] Wang T, Brede M, Ianni A, Mentzakis E. Detecting and characterizing eating-disorder communities on social media. Tenth ACM International Conference on Web Search and Data Mining - WSDM ’17; Feb 6-10, 2017; Cambridge, UK. 2017. pp. 91–100.

[13] Wang X, Zhang C, Sun L. An improved model for depression detection in micro-blog social network. IEEE 13th International Conference on Data Mining Workshops; Dec 7-10, 2013; Dallas, TX, USA. 2013. pp. 80–87.

[14] Working Party on Ethical Guidelines for Psychological Research. Code of human research ethics.Leicester, UK: The British Psychological Society; 2010.

[15] Brian A. Primack, César G. Escobar-Viera. Social Media as It Interfaces with Psychosocial Development and Mental Illness in Transitional Age Youth. 2017 , Child and adolescent psychiatric clinics of North America

[16] D.Sridhar, Dr.V.Kathiresan. The Psychology behind Users Mental Health in Social Media, International Journal of Computer Sciences and Engineering , May 2018