International Journal of Intelligent Information Systems
2021; 10(4): 69-73
http://www.sciencepublishinggroup.com/j/ijiis
doi: 10.11648/j.ijiis.20211004.15
ISSN: 2328-7675 (Print); ISSN: 2328-7683 (Online)

User Centric Social Opinion and Clinical Behavioural Model for Depression Detection

Ayodeji Olusegun Ibitoye*, Rantiola Fidelix Famutimi, Dauda Odunayo Olanloye, Ehisuoria Akioyamen

Computer Science Programme, Bowen University, Iwo, Nigeria

Email address:
Ibitoye_ayodeji@yahoo.com (A. O. Ibitoye)
*Corresponding author

To cite this article:
Ayodeji Olusegun Ibitoye, Rantiola Fidelix Famutimi, Dauda Odunayo Olanloye, Ehisuoria Akioyamen. User Centric Social Opinion and Clinical Behavioural Model for Depression Detection. *International Journal of Intelligent Information Systems*. Vol. 10, No. 4, 2021, pp. 69-73. doi: 10.11648/j.ijiis.20211004.15

Received: July 28, 2021; **Accepted**: August 20, 2021; **Published**: August 31, 2021

Abstract: In more recent time, depression as a lingering mental illness as continued to affect the way people act, and behave consciously or otherwise. Though it remained an undiagnosed disease globally without prejudice to age, gender, color or race; a lot of people never know implicitly or explicitly when they are depressed until it begins to affect their health conditions. While depression can be deciphered through text analysis in opinion mining, oftentimes, changes in human body also provides a convincing status of a depressed individual. No doubt, each data source can independently predict human depression status; however, the exclusive mutual relationship between both data sources has not been studied for depression detection. Therefore, in identifying meaningful correlations between clinical and behavioural data, this research detected depression by analyzing and matching mined patterns in users' behavioural opinion through tweets with trackable changes in clinical body vitals using wearable device for effective therapy in depressed patient management. Thus, by using a 5-fold cross validation on the clustered data, Random Forest ensemble model was used to build the Social-Health Depression Detection Model (SH2DM) after data preprocessing and optimal feature extraction. The dual data sourced user-centric model produced a better predictive result in accuracy, precision and recall values when compared and evaluated with single data depression detection instances of clinical and behavioural records.

Keywords: Depression Detection, Text-Analysis, Opinion Mining, Social-Health, Wearables, Random Forest, Decision Support

1. Introduction

In the race for sustainable and better life, everyone has the tendency to feel sometimes down or low, however, the cognitive functions of the human brain begin to depreciate if the emotional down times of an individual over a longer period remains steady. Several social factors like abuse, conflict, personal problem may lead to behavioural health disorders, depression as continued to be a pervasive problem among different age groups all over the world [1]. Depression denotes a wide range of mental health problems; it is pigeonholed by the absence of a positive vibes in a person's everyday activities. It is a serious medical condition that can get worse if not properly identified. It oftentimes observed in human expression and health condition. Consequently, identifying depression, is not dependent on just physical symptoms but also on behavioural, and social responses. Hence, to properly manage its severity, the need for the development of an automated system with capacity to monitor, detect and recommend appropriate therapy for depressed individuals became essential [2].

Over time existing methods such as self-report in depression detection analysis uses paper and/or blood testing method for depression screening. While the paper method is a multiple-choice questionnaire, which score obtained from the patients response determines the depression level, Blood tests or other extensive laboratory tests are used to make a conclusive diagnosis of the depressed patient [3]. In more recent times, monitoring and analyzing depression using a wearable sensor with machine learning has continued to provide meaningful insight into getting accurate results on depression implementation [4]. With the social media such as Twitter, Facebook, Instagram has become an unrestricted platform for free expression, machine learning algorithms are now been used to detect probable depressed user based on user social network and textual-analysis [5].

While there exist a connection between feelings and how such is expressed, it is difficult to categorically ascertain if both action and reaction match at same or difference instances. Basically, many times, human react (expressed opinion) in; or otherwise in contrast to how they feel. Also, the change in body chemistry varies among individuals and this requires special observation to ascertain once state of mind. Therefore, by using Random Forest ensemble model, this research employed learning, monitoring and matching patient's body vitals with expressed opinion in depression detection for operative therapy. The essence is to harness the power of behavioural and clinical symptoms in patience towards determining depression severity stages for appropriate decision support. Here, users public expressions are captured, and they do not need to take body vital readings manually. By this provision, a dual depression detection data centered model is implemented using IF-Association rules for better clustering of contributing depression vitals for a more actionable outcome. Thus, in section two, related works on the subject are presented, while the proposed depression detection model is presented in section three. In section four (4) the experiments and obtained results is discussed just as the research is concluded in section five.

2. Sample Related Works on Depression Detection

Due to denial or lack of routine diagnoses among others, in today's world, depression is being considered as one of the most common mental illness [6], which has been predicted to be the top contributor to Global Burden of Disease (GBD) by 2030 [7]. Although, it can be prevented by early symptoms detection, two out of three people suffering with depression do not enthusiastically seek for proper treatment [8]. While the patient mood is always affected first and can persist for a longer period, some of its social symptoms by [9] include avoidance of family and friends, loose of appetite, weight loss, physical illness among others. No doubt, by understanding the traits of depressed individual, early prediction of depression risk level can be diagnosed for efficient therapy. Although, the availability of specialized care for depressed patients is low, time, among other constraints from primary health care physician can also contribute to insufficient or delayed treatment [10, 11] for depressed patients. While it is not possible to present all related views on depression detection models in just an article, [12] investigation showed that the emotional and financial burden associated with depression can be significantly reduced if detection and appropriate intervention is provided at the early stage. To this end, there are lots of features that are needed to detect depression clinically and emotional. Consequently, several machine learning method have to be developed to analyze the collective data. Some of the works are related to behavioural analysis through tweets and others on clinical data has presented subsequently in section three (3).

Thus, towards enhanced and close monitoring of patience health status, in this modern era, wearable devices, which is a small sensors that is attached to a person's body has been widely adopted [13] in electronic health management. The essence is to cluster varieties of information that include but not limited to heart rate, blood pressure, amount of sleep, galvanic skin response and others [14]. For instance, by using a wearable textile device, [15] reported an accuracy of over 90% when predicting bipolar disorder in patients' emotional states from heart rate analysis. Also, [16] analyzed mood disorders by finding the correlation between objective features and depressive mood symptoms in patients using data collected from wearable devices. [17]. reviewed that the use of wearables helps to improve the safety and efficiency of intraoperative processes in clinical and simulated surgery.

In addition, aside clinical data for depression management, by analyzing opinion expressed on the social media e.g. Facebook, Twitter; identifying depressed individuals has also been made possible. For instance, [18] discovered that posting contents with negative emotional sentiments are common with depressed Twitter users. By this, several research techniques and approach have been developed over time. Some of this in review is [19] where predictive features with supervise learning algorithms was used to determine depressed and/or healthy contents. While [20] utilized the bag-of-words approach on a document level for better depression detection on a crowed source tweets, [21] considered sentiment analysis as a feature to detect depression from Twitter data. Similarly, [22] proposed a depression detection model based on 10 features and 3 classifiers to verify the model on Twitter data while a lexicon of terms was built [23] to identify depression or its symptoms from online opinion mining. With this and more, in section 3, the proposed depression detection model using clinical and behavioural data is presented for consideration. In section four, sample experiments and results are presented for evaluation.

3. Social-Health Depression Detection Predictive Model

The developed Social=Health Depression Detection Model (SH2DM) detects and clusters periodic clinical data using wearable devices. The clinical data is constantly transmitted through the use of Bluetooth to a friendly built Android based mobile application for depression management. At the first instance, user's preliminary biodata are required and this also includes unique twitter handle e.g. @deji22 while tweets activities are closely monitored using the Twitter API. The body vitals as captured from the wearable devices, include heartbeat, body temperature, number of steps, and sleep time. The data capturing activities (clinical and opinion data) are periodically clustered on both depressed patients and healthy volunteers for analysis during the period of the study. The essence is to predict without restriction users depression status in real time irrespective of the kind of activities they are engaged. Thus from the clustered data, other basic machine learning actives on the data are performed as presented in figure 1.

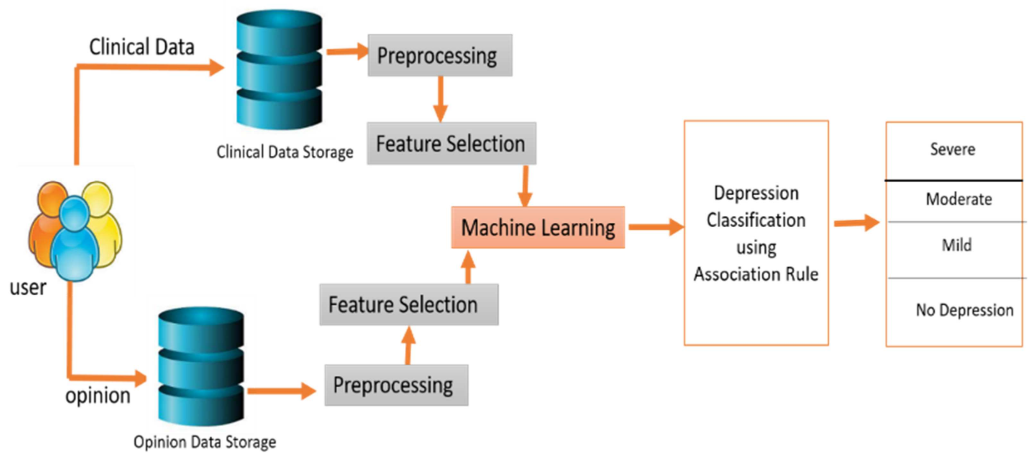

Figure 1. *Social-Health Depression Detection Model.*

Therefore, from figure 1, clinical and opinion data are uniquely stored and preprocessed for machine learning. First, data tracked from the wearable devices was preprocessed hourly. Then the standard deviation values are computed per day to create unique features for use in machine learning models over a period of 90 days. For the text preprocessing, the Natural Language Preprocessing Tool Kit (NLPTK) was used for tokenization, parsing, and sentiment extraction. The optimised feature subsets were used for building the machine learning models, which engenders the depression categories. Then, the If-Association rules were used to match corresponding classification of outputs from clinical and opinion dataset based on available criteria. It is important to state that, depression prediction through text analysis does not process distinctly the tweets of users for the user's classification. However, output from the collective tweet analysis of the user community as a prevailing impact on the depression status of the concerned. Therefore in this experiment, users depression categories were predicted as Severe (0.75 – 1), Moderate (0.50 – 0.74), Mild (0.25 – 0.49) and No Depression (0 – 0.24). For the learning algorithm, 5-fold cross-validation was applied for the validation on the Random Forest machine learning model. In section five (5), sample experiments and evaluations are presented.

4. Experiments and Evaluations

From a clustered 356 participants in a period of 3 months, without the inclusion of personal information that can reveal an identity, the scope of the work experimented a dual data sourced and data preprocessing machine learning model for depression detection. 34 distinct features were engenders from users wearable devices while a total of 3025 unique tweets were clustered based on users activities. The experimented data representation is presented in table 1.

Table 1. *Data representation.*

Activities	Control	Depressed
Total users	197	159
Total Number of Tweets	2076	949
Average Age	22.7	20.5
Gender (Female) Distribution	63%	76%

Although, data imbalance was a constraint on the experimental dataset, Accuracy, Precision, and Recall, were standard metrics used for evaluation in this research. When users are inaccurately predicted, a low precision value is obtained. i.e., False Positives (FP). However, when users who are depressed are not recognised, a low recall i.e. False Negative (FN) result are obtained. Thus, the acquired results after analysis are presented as follows:

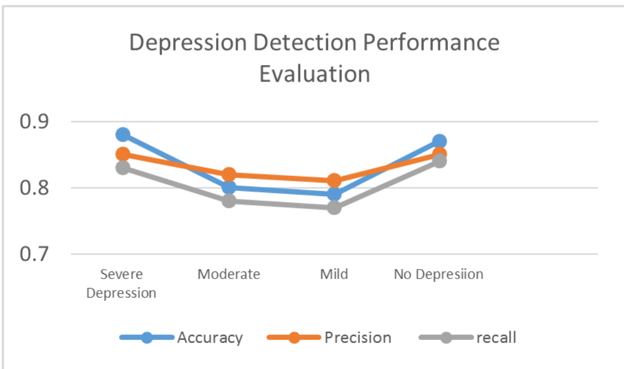

Figure 2. *Performance Evaluation.*

Based on just two depression classification types (depressed and non-depressed) the developed model was also evaluated distinctly with wearable devices and textual analysis as an entity. The following results as presented in table 2 were obtained.

Table 2. *Model Performance Evaluation.*

Cat	Accuracy			Precision			Recall		
	SH	**W**	**O**	**S**	**W**	**O**	**S**	**W**	**O**
SD	0.88	0.79	0.76	0.85	0.82	0.81	0.83	0.78	0.76
ND	0.87	0.75	0.70	0.85	0.79	0.76	0.84	0.81	0.78

SH: the proposed model, W- Wearable, O – Opinion mining.

From the available result, we observed that the developed SH2DM model, outperformed its single entity approach using random forest ensemble classification model. Although, the small number of participants in this research, may not represent the true quality of statistical importance, yet we have observed that a quadruple depression detection classification is achievable with an imbalance data from two different data sources (clinical and textual). The combination of the two mutually exclusive data source is structured to measure the influence of a user's social community and its impact on body vitals towards predicting human depression status and vice versa.

5. Conclusion

The quality of data and the type of processing, no doubt, impacts an outcome. In the article clinical and textual data were clustered to predict the depression status of human. Through text analysis in a user community, the depression status of user is flagged and matched with preprocessed clinical data. From data preprocessing to machine learning, an if Association rule was used to classify depression status into four-class category. From all experiments and evaluation, the developed SH2DM model performed better and will be used in the nearest future on larger dataset as we also like to investigate the depression detection management on global factors like Gender, Colour and Ethnicity.

References

[1] Gitanjali S, Iachan R, Scheidt PC, Overpeck MD, Sun W, Giedd JN: Prevalence of and Risk Factors for Depressive Symptoms Among Young Adolescents. Arch Pediatr Adolesc Med 2004, 158: 760-65.

[2] Maurer DM, Raymond TJ, Davis BN. Depression: Screening and diagnosis. Am Family Phys (2018) 98: 508–15.

[3] Gelaye B, Williams MA, Lemma S, Deyessa N, Bahretibeb Y, Shibre T, Wondimagegn D, Lemenhe A, Fann JR, Vander Stoep A, et al. Validity of the Patient Health Questionnaire-9 for depression screening and diagnosis in East Africa. Psychiatry Res. 2013; 210 (2): 653–61.

[4] Saad M., Ray L. B., Bujaki B., Parvaresh A., Palamarchuk I., De Koninck J., Douglass A., Lee E. K., Soucy L. J., Fogel S., Morin C. M., Bastien C., Merali Z., Robillard R. Using heart rate profiles during sleep as a biomarker of depression. BMC Psychiatr. 2019; 19: 168.

[5] De Choudhury, M., Gamon, M., Counts, S. and Horvitz, E., 2013, June. Predicting depression via social media. In Seventh international AAAI conference on weblogs and social media.

[6] Yazdavar, A. H., Al-Olimat, H. S., Banerjee, T., Thirunarayan, K., & Sheth, A. P. Analyzing clinical depressive symptoms in twitter (2016).

[7] Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. PLoS medicine, 3 (11), e442.

[8] "Media- DBSA Facts about Depression-, Secure2.convio.net" 1 January, 2019. [Online]. Available: https://secure2.convio.net/dabsa/site/SPageServer/;jsessionid=00000000.app20101a?NONCE_TOKEN=B2DB 3A070406934661E7561D. [Accessed 21 November 2019].

[9] Timothy J. Legg, "The effects of depression on the body and physical health. [online]," Medical News Today, 1 January 2019. [Online]. Available: https://www.medicalnewstoday.com/articles/322395.ph p. [Accessed 21 November 2019].

[10] Unützer J, Park M. Strategies to improve the management of depression in primary care. Prim Care. (2012) 39: 415–31. doi: 10.1016/j.pop.2012.03.010.

[11] Lin EH, Katon WJ, Simon GE, Von Korff M, Bush TM, Walker EA, et al. Low-intensity treatment of depression in primary care: is it problematic? Gen Hosp Psychiatry. (2000) 22: 78–83. doi: 10.1016/S0163-8343(00)00054-2.

[12] Halfin A. Depression: the benefits of early and appropriate treatment. Am J Manag Care 2007 Nov; 13 (4 Suppl): S92-S97.

[13] Mazzetta I., Gentile P., Pessione M., Suppa A., Zampogna A., Bianchini E., Irrera F., Stand-alone wearable system for ubiquitous real-time monitoring of muscle activation potentials. Sensors. 2018; 18: 1748.

[14] Nieto-Riveiro L., Groba B., Miranda M. C., Concheiro P., Pazos A., Pousada T., Pereira J. Technologies for participatory medicine and health promotion in the elderly population. Medicine (Baltim.) 2018.

[15] Valenza G., Nardelli M., Lanatà A., Gentili C., Bertschy G., Paradiso R., Scilingo E. P. Wearable monitoring for mood recognition in bipolar disorder based on history-dependent long-term heart rate variability analysis. IEEE J. Biomed. Health Inf. 2014; 18: 1625–1635.

[16] Rohani D. A., Faurholt-Jepsen M., Kessing L. V., Bardram J. E. Correlations between objective behavioural features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: systematic review. JMIR mHealth uHealth. 2018.

[17] Lauren Kolodzey, Peter D Grantcharov, Homero Rivas, Marlies P Schijven, Teodor P Grantcharov, Wearable technology in the operating room: a systematic review on behalf of the Wearable Technology in Healthcare Society.

[18] Park, G., et al. Automatic personality assessment through social media language. J. of Pers. and Soc. Psychol. 108 (6), 934 (2015).

[19] Reece, A. G., et al. Forecasting the onset and course of mental illness with Twitter data. Sci. Rep. 7 (1), 13006 (2017).

[20] Nadeem, M. Identifying depression on Twitter. Preprint at arXiv: 1607.07384 (2016).

[21] Mowery, D., Bryan, C., & Conway, M. Feature studies to inform the classification of depressive symptoms from Twitter data for population health. Preprint at arXiv: 1701.08229 (2017).

[22] Wang, X., Zhang, C., Ji, Y., Sun, L., Wu, L. & Bao, Z. (2013). A depression detection model based on sentiment analysis in micro-blog social network. In PacificAsia Conference on Knowledge Discovery and Data Mining (pp. 201-213). Springer, Berlin, Heidelberg.

[23] De Choudhury, M., Counts, S. and Horvitz, E., 2013, May. Social media as a measurement tool of depression in populations. In Proceedings of the 5th Annual ACM Web Science Conference (pp. 47-56). ACM.