A TWO-LEVEL DETECTION AND RECOVERY RECOMMENDATION SYSTEM FOR DEPRESSION

Ayuba Philip S.[1*], Dr. O. A. Ojerinde[2], Dr. H. O. Aliyu[3]
[1]Department of Computer Science, Federal University of Technology Minna, Nigeria
[2] Department of Computer Science, Federal University of Technology Minna, Nigeria
[3] Department of Computer Science, Federal University of Technology Minna, Nigeria
*Corresponding author: philipayuba16@gmail.com, +234-7061251979

Abstract
Depression is an emerging problem in public health. Various socio-demographic factors like age, sex, earning status, living spouse and family type, etc. are responsible for depression among people. Some co-morbid conditions like visual problems, hearing difficulties, and mobility problems also influence the disease. But depression can be diagnosed at the earliest using predictive modeling with various influencing input variables. The paper focuses on developing a system that can identify depression on multiple levels and provide a model for recovery. Waikato Environment for Knowledge Analysis (WEKA), a data mining tool was used for prediction based on machine learning classifiers. Also, five machine learning classifiers were compared with respect to three test options where extracted facial numerical attributes were iterated twice using a Multilayer Perceptron (MLP) classifier, with an accuracy of 82% and 89%, and precision of 87.2% and 89.3%. The system has the potential to lower depression statistics by employing a more efficient detection algorithm that takes into account both visual and text inputs including a well-designed rehabilitation model.

Keywords:

Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Depressions, Waikato Environment for Knowledge Analysis (WEKA),

1.0 INTRODUCTION

Depression is a common mood disorder which affects the lives of the individual suffering from it. It is a worldwide problem affecting innumerable lives. Depressed people are more prone to anxiety, sadness, loneliness, hopelessness, disruption in emotion experience, communication loss, lack of self-regulation and are frequently worried and disinterested (Dham et al., 2017; Yang et al., 2013). According to the Diagnostic and Statistical Manual of Mental Disorders, the most widely used resource in diagnosing mental disorders in the United States, most people will experience some feelings of depression in their life-time, although it does not meet the criteria of an illness until a person has experienced, for longer than a two-week period, a depressed mood and/or a markedly diminished interest/pleasure in combination with four or more of the following symptoms: significant unintentional weight loss or gain, insomnia or sleeping too much, agitation or psychomotor retardation noticed by others, fatigue or loss of energy, feelings of worthlessness or excessive guilt, diminished ability to think or concentrate, indecisiveness, or recurrent thoughts of death (Morales et al., 2017).

Given advancements in hardware and software, coupled with the explosion of smartphone use, the forms of potential health care solutions have begun to change and interest in developing technologies to assess mental health has grown. Among the latest technologies are depression detection systems, which use indicators from an individual in combination with machine learning to make automated depression level assessments. This puts the ability to detect depression at a more accurate level in your hands, giving us the opportunity to take better care of our mental health (Morales et al., 2017).

Visual indicators have been widely explored for depression analysis, including body movements, gestures, subtle expressions, and periodical muscular movements (Morales et al., 2017). Facial expressions play an important role in the communication of emotion. Features are usually grouped in two types of facial descriptors: appearance and geometric based (Valstar et al., 2016).

Automatic detection of depression has attracted increasing attention from researchers in psychology, computer science, linguistics, and related disciplines. As a result, promising depression detection systems have been reported (Morales et al., 2017). However, it is not just enough to detect depression anymore there must be provision of some immediate help in recovery that is particular to the subject to help slow down the depression process and prolong the individuals life (National Institutes of Health, 2007).

According to the World Health Organization, depression is the most common mental disorder. Currently, 300 million people suffer from depression (WHO) (Havigerová et al., 2019). According to the Federal Ministry of Health, 30% of Nigerians suffer from mental illness and with a population of about 200 million, Nigeria has a high rate of mental illness. This implies that Nigeria has about 60 million persons with mental illnesses. Nigeria, in essence, has more mental cases than the population of every West African country (Sahara, 2018). In 2017, the World Health Organization stated that 7,079,815 Nigerians suffered from one of the most ignored and misunderstood forms of mental disorder in the country is depression. The purpose of this study is to reduce these statistics drastically using a more efficient detection technique that would consider both visual and text input and design a recovery model.

2.0 Literature Review

In 2021, Gutierrez concluded that the internet of things IoT can help mental illness patients by building well-being systems with a particular focus on prominent aspects such as data acquisition, self-organization, service level agreement, security, and identity management (Gutierrez et al., 2021).

(Katherine et al, 2017) concluded that recovery from depression is perceived by patients as a complex, personal process that is influenced by a range of factors. However, a greater understanding of clinicians' perceptions of client recovery from depression would be beneficial to inform clinical practice and influence future research.

(Wang et al., 2018) developed a system for detecting depression for students known as the student life app. As shown in figure 2.0, the StudentLife sensing app is built based on our prior sensing work for Android and iOS phones. The StudentLife app continuously infers and records participants' physical activities (e.g., stationary, in a vehicle, walking, running, cycling), sleep (duration, bed time, and rise time) based on a prior work on sleep inference using phones, and sociability (that is, the number of independent conversations a participant is around and their duration). The app also collects audio amplitude, location coordinates, and phone lock/unlock events. A built-in MobileEMA component is used to administer self-reported. The app uploads the data to the secured server when the user is charging their phones and under WiFi. StudentLife is extended to collect data from wearables; specifically, physiological data was collected from Microsoft Band 2 and given to each of the students in the study. The StudentLife app collects the heart rate, galvanic skin response (GSR), skin temperature, and activity data from the band in real-time. Band data is uploaded to the StudentLife app over Bluetooth and then uploaded to the servers as described above. Note, during data modelling and analysis poor data quality issues associated with GSR and skin temperature data from the band were found. First, the GSR sample rate provided by the Microsoft Band SDK is too low to be useful in analysis. Such a low sample rate limited extracting of useful GSR features. It was also discovered that the skin temperature sensor reading is affected by the ambient environment temperature. The temperature differences between indoor and outdoor during New Hampshire winter can be as large as 70 degrees Fahrenheit. It was observed significant drops in skin temperature when participants are outside during the winter term. For these reasons, GSR and skin temperature data in the modelling was not used. While the StudentLife app infers sleep data from the phone, only within +/- 25 minutes of error, the band has much better sleep measurements. However, because the band only lasted one day due to limited battery and the demands of continuous sensing most students wore the band during the day and recharged it at night. The result had limited sleep data from the band. Therefore, only sleep measurements from the phone data was used. Even though there was a collection of GSR, sleep, skin temperature the researcher ended up using only heart rate and activity data from the bands (Wang et al., 2018).

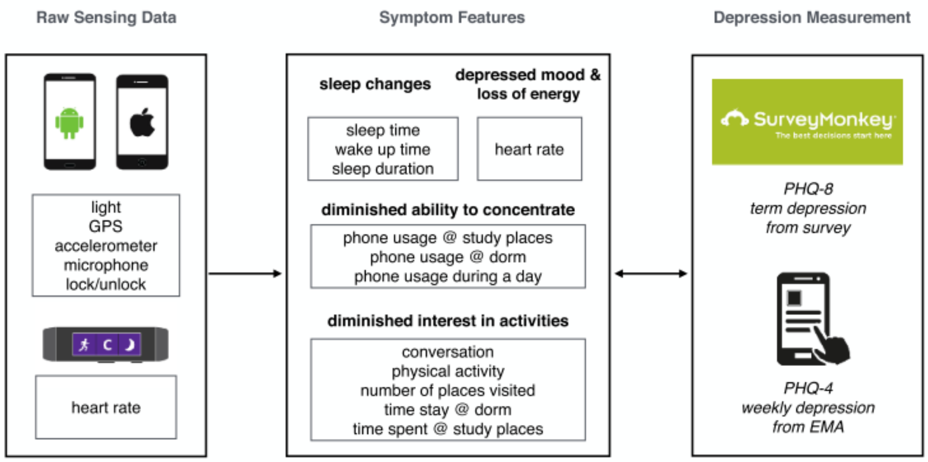

Figure 2.0: Wrist Band for sensing depression. (Wang et al., 2018)

3.0 Methodology

3.1 Data collection

As the raw data have a certain degree of noisy and redundant information which is irrelevant to depressive expressions. To extract meaningful information from noise, it is necessary to apply multiple pre-processing steps on the raw data before feeding it to the model. To deal with the raw data, face detection and landmark localization of each subject in the video are implemented in Dlib (Johnston et al, 2018). After that, the facial region of an image size of 256×256 is cropped and aligned according to the eye locations.

3.2 Data Preprocessing

Conditions during data collection are not always conveniently established, and therefore preprocessing can help in obtaining meaningful information for further analysis. The relevant algorithms employed during different experimental tests of this paper included:

a) Illumination normalization

b) Face detection

c) Facial landmarks detection

d) Face alignment.

The specific algorithms for each preprocessing step are explained within the following subsections.

3.3 Sample Space

Eligible participants were 18 years and older, they either have a first degree or enrolled in an institution in order to get one. A simple Google form would be used to get information and bio data for participants to be sure they fit the criteria, participants are also offered an incentive of a chance to win a ₦2000 airtime voucher in a raffle draw. Their details will be kept confidential and used for this purpose only.

3.4 Instruments and Measures

Participants were asked the following questions to determine their level of depression via text (Longnecker, 2017; Wang et al., 2018), Questions would be asked on a five-point scale:

1. Little interest or pleasure in doing things?
2. Feeling down or helpless?
3. Trouble falling asleep or sleeping too much?
4. Feeling tired
5. Not eating or over eating
6. Trouble concentrating on things?
7. Thoughts of hurting yourself, or that you could be better off none existent.

3.5 Flow Diagram for the new depression detection and recovery system

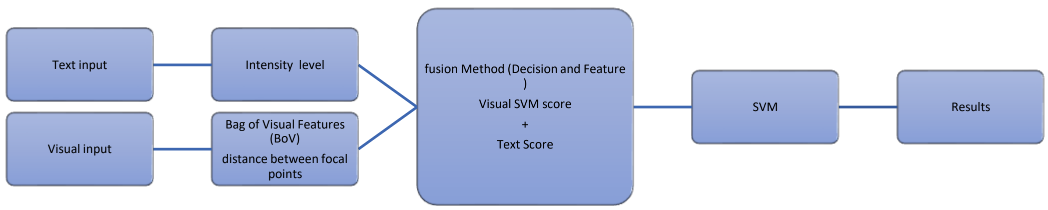

Figure 3.5: Depression Detection Part of the System

3.5.1 Feature Fusion

This is the simplest form of fusion. Raw features computed from the different modalities are concatenated to form a single feature vector. Despite the simplicity, feature fusion results in a performance increase compared to the performance of single modalities. However, the downside of feature fusion is that it suffers from the curse of dimensionality (Joshi et al., 2013).

3.5.2 Score Fusion

In score-level fusion, different scores such as probability estimates, likelihoods, etc. are combined, before making a classification decision. There are several popular methods for score fusion. In this paper, two techniques – score fusion by weighted sum and by learning a new SVM classifier on the scores – are investigated. The distance from the SVM hyperplane is calculated and used as a score (Joshi et al., 2013)

3.5.3 Decision Fusion

In decision fusion, multiple classifiers are trained on different feature sets. The output of these classifiers is used to infer the final class result. In this paper, the AND and OR operators are used to fuse the decisions from the separate text and visual SVM classifiers.

3.6 Flow Chart Diagram of the Proposed System

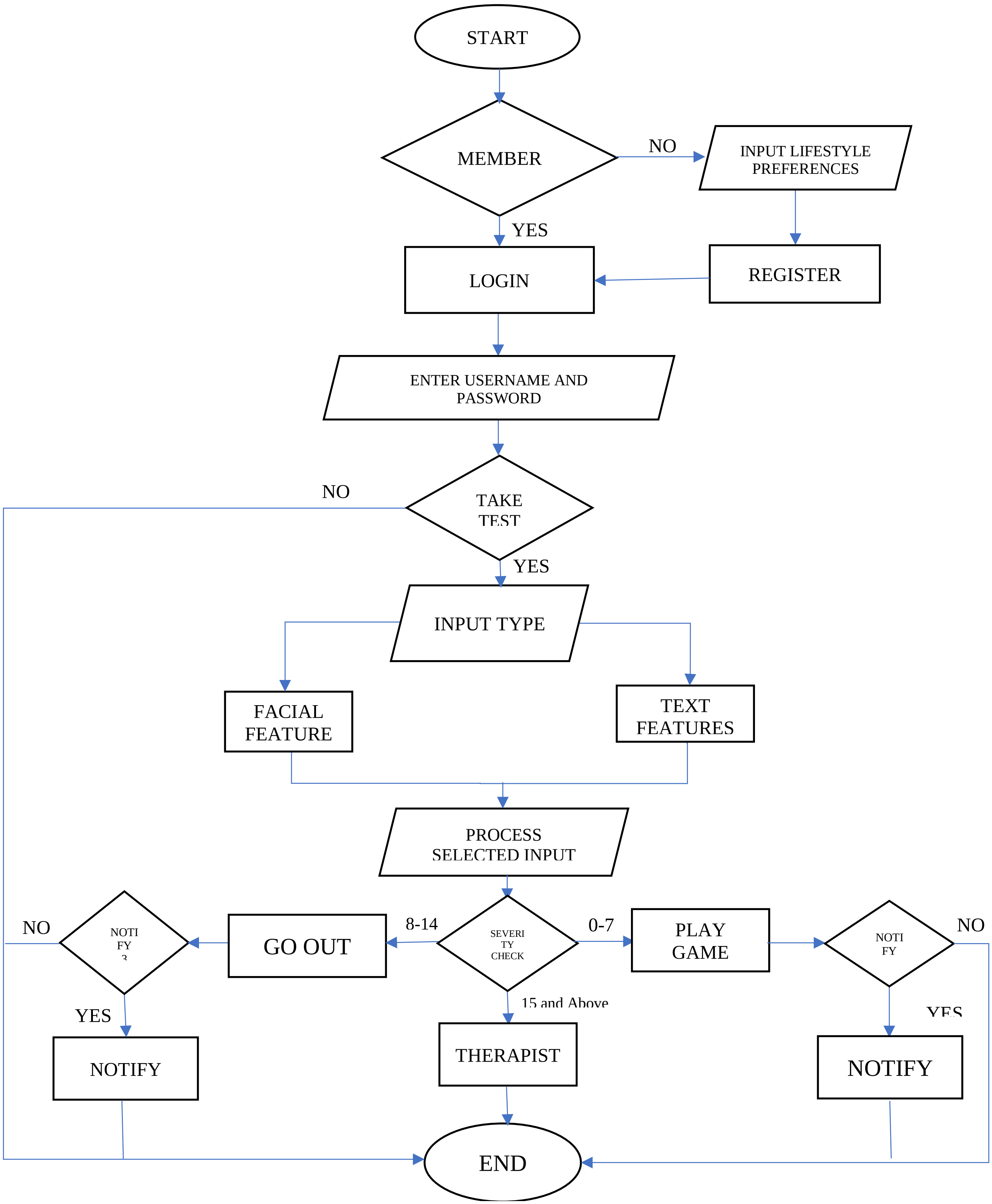

Figure 3.6: Flow Chart Diagram of the Proposed System

The user in the above flowchart diagram while using the system is initially prompted to login as a new member or an old member who already has an account. A new member is obligated to register his details which include name, date of birth, address, phone number, health records and most especially his/her lifestyle preferences. After the registration process, an assigned username and password are used to login and the depression detection test is carried out. A user who does not intend to carry out the test is taken to the end page of the system.

For users who take the test, inputs of Facial and Text features are collected and processed. A result is generated in line with the Hamilton Rating Scale of Depression thus providing the user's level of depression.

Suggestive measures are provided to the user to recover from depression depending on the Level of depression gotten from the above test results. A constant notification of 3 days interval is a major feature to help the user keep track of his/her progress toward recovery. If the case is severe, the user is forwarded to a therapist for in-depth analysis and treatment.

3.7 Machine Learning

The machine learning level is concerned with exploiting the extracted features in order to train models that generalize well and are able to recognize signs of depression. Given the high dimensionality of the features, there is a need to employ a reduction algorithm. Further, before the training/testing of the models there is another issue to address, this of the cross-validation which is used so as to deal with less biased measures. Finally, in the proposed work two approaches were investigated, the categorical (classification) and the continuous (regression).

3.8 Measuring depressive behaviour

A set of attributes like emotional process, temporal process, and linguistic style can be used to characterize the depressive behaviours of users. Our dataset consists of five emotional variables (positive, negative, sad, anger, anxiety), three temporal categories (present focus, past focus and future focus), and 9 standard linguistic dimensions (for example articles, prepositions, auxiliary verb, adverbs, conjunctions, pronoun, verbs and negations)

3.9 Convolutional Neural Network (CNN)

Convolutional neural network (CNN, or ConvNet) is a class of artificial neural networks, most commonly applied to analyze visual imagery. As the term "convolutional neural network" suggests, this system makes use of convolutional mathematical operations. At least one of the layers of convolutional neural networks uses convolution instead of standard matrix multiplication (Goodfellow, 2016).The ability to extract features without the need for prior knowledge or human intervention is a major advantage of CNN.

3.9.1 Applications of CNN

Some of the applications of CNN are Image recognition, video analysis, natural language processing, anomaly detection, drug discovery, health risk assessment of biomarkers of ageing discovery, checker games, time series forecasting,

3.10 Data Analysis and Depression System for Text-based System

ANN predictive modelling, its performance analysis and development of a depression prediction system was done using WEKA version 3.8.0 software. Weka is an open-source data mining software in the field of machine learning technology. The predictive model was built using ANN (Multilayer Perceptron, A feed-forward ANN Model) as a classifier in the Weka. The performance of the build model was automatically analysed using a 10-fold cross-validation method.

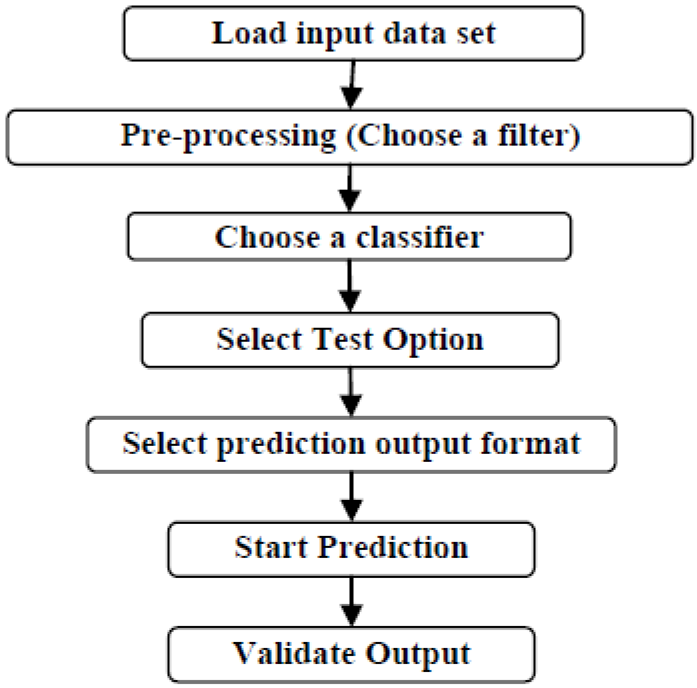

Figure 3.10: Predicting depression in WEKA

3.10.1 Loading ANN Model

Multilayer Perceptron (MLP), a feed-forward ANN model was built in WEKA for classification with input variables i.e., Little interest or pleasure in doing things? Feeling down or helpless? Trouble falling asleep or sleeping too much? Feeling tired, Not eating or overeating, Trouble concentrating on things?

Thoughts of hurting yourself, or that you could be better off none existent. The ANN model was designed with seven interconnected neurons in one hidden layer. The learning rate of the network was set at 0.3 for 50 iterations. Output was classified as normal i.e., non-depressed and depressed. For the facial dataset, CNN's feature extraction techniques was used to extract numerical features from the images and fed them into the Classification system.

3.10.2 Pre-processing

After loading the input data set, filter is applied from a preprocessing window. A supervised filter – attribute Selection is applied in this research work to select optimized attributes.

3.10.3 Classification

After applying the filter, input data set is fed into classify window. After that classifiers are chosen for classification. Here five different classifiers are chosen and run the experiments five times and classification output is collected each time.

RESULTS AND DISCUSSION

4.1 Depression Data collection

Facial datasets were gotten from (*Kaggle Repository*, 2013), a Kaggle repository that is divided into train (80%), test (10%), and validation (10%).

4.2 Experimental setting

Numerical features are extracted from the images into textual representation using CNN and fed into the classification system. These features are pre-processed to remove outliers, nulls, and invalid data from the feature set.

4.3 Paradigm Design of Depression Experiments

Personal experiences, facial expressions, individual differences in nervous system response, and emotional responses are all associated with depressed individuals' negative self-schema in cognitive processing related to attention-control disorder. Using the methodology provided, the depression experiment tasks are divided as follows:

1. **Textual quiz:** 7 short textual quizzes, with 5 scales of answers were presented. Participants were required to answer those quiz questions on a web-based platform. The questions are excerpted from. The quiz questions are:

a. Little interest or pleasure in doing things?

b. Feeling down or helpless?

c. Trouble falling asleep or sleeping too much?

d. Feeling tired

e. Not eating or overeating

f. Trouble concentrating on things?

g. Thoughts of hurting yourself, or that you could be better off none existent.

2. **Facial input:** The user is prompted to a webcam facial input in which his/her face data is gotten and fed into the neural network.

4.4 CONSTRUCTING DEPRESSION DATASET

Each user has to complete any of the two tasks to form their own dataset for classification. The facial input is recorded in .jpg format and the textual quizzes response are in .csv format. The information relating to the subject is saved in the information log, which includes name, age, gender, educational background, profession and so on.

4.5 EXPERIMENTAL RESULTS

Weka 3.8.0 was used for the experiment. WEKA is a data mining tool used for prediction based on machine learning classifiers. Comparison between depressed and non-depressed machine learning data with various sociodemographic factors and their weights as shown in Table 2.

Table 1:Experimental Results

Variables	**Strongly Agree**	**Agree**	**Maybe**	**Disagree**	**Strong Disagree**
Little interest or pleasure in doing things?					
Depressed	5	7	3	1	1

Non-Depressed	1	1	3	5	7
Feeling down or helpless?					
Depressed	5	7	3	1	1
Non-Depressed	1	1	3	5	7
Trouble falling asleep or sleeping too much?					
Depressed	5	7	3	1	1
Non-Depressed	1	1	3	5	7
Feeling tired					
Depressed	5	7	3	1	1
Non-Depressed	1	3	3	5	7
Not eating or overeating					
Depressed	5	7	3	1	1
Non-Depressed	3	2	3	5	5
Trouble concentrating on things?					
Depressed	5	7	3	1	1
Non-Depressed	3	1	3	5	7
Thoughts of hurting yourself, or that you could be better off none existent.					
Depressed	5	7	3	1	1
Non-Depressed	1	1	3	5	7

4.6 PERFORMANCE MEASUREMENT

The preprocessed data were fed into the system and iterated twice. The results are presented in Table 2. The extracted facial numerical attributes were iterated 2 times using MLP classifier, with an accuracy of 82 and 89 %, and precision is 87.2 and 89.3. ROC area (Area Under the Curve) is 0.92 and 0.95.

Table 2: Result of prediction with Cross-validation

Type	**Metric**	**Accuracy (%)**	**Precision**	**ROC Area**	**RMSE**
Facial	MLP	82	87.2	0.92	0.33
Facial	MLP	89	89.3	0.95	0.29
Textual Quiz	MLP	83	85.1	0.96	0.31
Textual Quiz	MLP	97	92.5	0.93	0.31

4.6.1 INTERPRETATION OF METRICS

Accuracy: Accuracy is the proportion of correctly classified instances out of the total number of instances. A high accuracy indicates that the model is able to correctly classify most of the instances. On the textual metrics, 83% and 97% perform better on average than the 82% and 89% average of the facial classification.

Precision: Precision is the proportion of correctly classified instances in a particular class (i.e., true positives) out of the total number of instances classified as belonging to that class (i.e., true positives and false positives). A high precision indicates that the model is able to correctly identify most of the instances in that class, and there are few false positives. The range of 87.2 to 92.5 records a high precision.

ROC Area: ROC (Receiver Operating Characteristic) Area is a measure of the performance of a binary classifier. It represents the area under the ROC curve, which plots the true positive rate (sensitivity) against the false positive rate (1 - specificity) at different classification thresholds. A high ROC Area indicates that the model is able to correctly classify most of the instances while keeping the false positive rate low. ROC Area is noted to be higher for the two classifications.

RMSE: RMSE (Root Mean Squared Error) is a measure of the difference between the predicted and actual values in a regression problem. A low RMSE indicates that the model's predictions are close to the actual values.

CONCLUSION

Having depression may have a significant impact on one's mental health and everyday routine. Recently, researchers have been seeking an objective assessment approach and quantitative markers that can accurately detect depression. Research into the detection of depression based on facial expressions is a trendy issue among recent studies. In this study, a two-level Detection and recovery recommendation systems for Depression was designed effectively stimulating the use textual and facial inputs to establish depression levels in individuals. Two network models that collaborate with each other have been used. While the first is based on literature (Longnecker, 2017; Wang et al., 2018), to determine the level of depression on a 5-scale answer, the other uses CNNs to analyze visual inputs of the users and compare against the already trained dataset.

REFERENCE

Dham, S., Sharma, A., & Dhall, A. (2017). *Depression Scale Recognition from Audio, Visual and Text Analysis*.

Gutierrez, L. J., Rabbani, K., Ajayi, O. J., Gebresilassie, S. K., Rafferty, J., Castro, L. A., & Banos, O. (2021). Internet of things for mental health: Open issues in data acquisition, self-organization, service level agreement, and identity management. *International Journal of Environmental Research and Public Health*, *18*(3), 1–19. https://doi.org/10.3390/ijerph18031327

Havigerová, J. M., Haviger, J., Kučera, D., & Hoffmannová, P. (2019). Text-based detection of the risk of depression. *Frontiers in Psychology*, *10*(MAR), 1–11. https://doi.org/10.3389/fpsyg.2019.00513

Johnston, B., & Chazal, P. de. (2018). A review of image-based automatic facial landmark identification techniques. *Eurasip Journal on Image and Video Processing*, *2018*(1). https://doi.org/10.1186/s13640-018-0324-4

Joshi, J., Goecke, R., Alghowinem, S., Dhall, A., Wagner, M., Epps, J., Parker, G., & Breakspear, M. (2013). Multimodal assistive technologies for depression diagnosis and monitoring. *Journal on Multimodal User Interfaces*, *7*(3), 217–228. https://doi.org/10.1007/s12193-013-0123-2

Kaggle Repository. (2013). https://www.kaggle.com/astraszab/facial-expression-dataset-image-folders-fer2013

Katherine Richardson. (2017). *This is a repository copy of Recovery from depression : a systematic review of perceptions and associated factors . White Rose Research Online URL for this paper : Version : Accepted Version Article : Richardson , K . and Barkham , M . orcid . org / 0000-*.

Longnecker, E. M. (2017). The Relationship between Smartphone Use, Symptoms of Depression, Symptoms of Anxiety, and Academic Performance in College Students. *ProQuest Dissertations and Theses*, 72.

Morales, M., Scherer, S., & Levitan, R. (2017). *A Cross-modal Review of Indicators for Depression Detection Systems*. 1–12. https://doi.org/10.18653/v1/w17-3101

National Institutes of Health. (2007). Information about Mental illness and the Brain. *NIH Curriculum Supplement Series*, *Md*, 1–35.

Sahara, R. (2018). *30% Of Nigerians Suffer Mental Illness – FG*. https://saharareporters.com/2018/11/13/30-nigerians-suffer-mental-illness-—-fg

Valstar, M., Gratch, J., Schuller, B., Ringevaly, F., Lalanne, D., Torres, M. T., Scherer, S., Stratou, G., Cowie, R., & Panticz, M. (2016). AVEC 2016 - Depression, mood, and emotion recognition workshop and challenge. *AVEC 2016 - Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge, Co-Located with ACM Multimedia 2016*, *October*, 3–10. https://doi.org/10.1145/2988257.2988258

Wang, R., Wang, W., daSilva, A., Huckins, J. F., Kelley, W. M., Heatherton, T. F., & Campbell, A. T. (2018). Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(1), 1–26. https://doi.org/10.1145/3191775

Yang, Y., Fairbairn, C., & Cohn, J. F. (2013). Detecting depression severity from vocal prosody. *IEEE Transactions on Affective Computing*, *4*(2), 142–150. https://doi.org/10.1109/T-AFFC.2012.38