



Binary Logistic Gradient Transformation Based Behavioral Modeling for the Early Mental Health Illness Diagnosis

Sandeepkumar Hegde¹, Rajalaxmi Hegde²

¹NMAM Institute of Technology, India, sandeep.hegdey@gmail.com

²NMAM Institute of Technology, India, laxmi123.prabhu@gmail.com

ABSTRACT

Mental health is a pointer of the passionate, mental and social well-being of a person. It decides how a person considers, feels and handle circumstances. Maintaining the proper mental healthiness is very important for the working professional to work productively with full potential in the working environment. Due to the extensive competition in the working environment, the working professional often suffers from the problem like stress, anxiety, depression. Hence these problems may lead to mental illness. If the mental illness is not diagnosed in the early stage, it may lead to a problem like suicides. The mental illness is considered as a chronic disease which will persist for a longer duration. It is difficult to diagnose such disease in advance by the Physician. Hence Machine learning is considered as one of remedy to early predict the mental illness. In the proposed paper Binary Logistic Gradient Transformation(BLGT) based Behavioral Modeling has been proposed for the early prediction of mental illness among working professionals. Mental health data is gathered from the UCI repository. The BLGT will choose the subset of the features from the data which maximize the likelihood and minimize the square error by calculating the logistic score. The model transformations are done on a distributed set of random variables using the sigmoid function to restrict the response to bounded outcome score which maps the probability values between the binary classes. The accuracy obtained through the proposed model is compared with various machine learning algorithms and also validated using the various parameters such as precision, recall, f measure, confusion matrix, and ROC. The experimental results illustrate that the proposed BLGT model is outperformed in accuracy compared to the existing approaches.

Key words : Behavioral Model, Machine Learning ,Mental health, , Model Transformation, Sigmoid function

1. INTRODUCTION

In this competitive era, the working professional always engaged in work assignments. Constantly involving in work assignment is good from the point of view of the mental healthiness. But if there is a negative working environment, it may lead to mental health problems. The effect of the mental health issues within the work environment not only impacts

the efficiency of a particular individual but also bring down the productivity of the entire enterprise. It further leads to performance issues, absenteeism, accidents, alcoholism, lost productivity. The mental disorder like anxiety and depression are leading cause or threat which restrict the employability of individuals[1]. Hence it negatively affects the self-esteem, income, and quality of life. Approximately one in 100 people will be affected by a mental disorder at some point in their lives and will struggle with them for decades. Hence the early diagnosis of mental health illness is very important among working professionals. In this paper Binary Logistic Gradient Transformation(BLGT) based Behavioral Model has been proposed for the early prediction of mental illness among working professionals. The dataset is accumulated from UCI[2] gathered contained details about 1260 working professional. Initially model will select the optimal features from the input data using Symmetric Disparity Based Feature Selection Algorithm(SDFS). The SDFS will select the features from the input dataset which will have a direct impact on the class label[3]. The proposed BLGT model trained on the mental health data set containing the features extracted from the feature selection algorithm. The model transformation is done through the complex sigmoid function which will map the probabilistic prediction value between binary classes[4]. In proposed model Cross Entropy-based cost functions are adapted which will induce the penalty whenever the predicted probability differs from the actual label. The error rate of the constructed model is minimized by using gradient descent function. It is a local optima function which will update the model coefficient automatically. The 10 fold cross-validation is performed on the whole model to overcome the problems like model overfitting and underfitting. The implemented model is validated using various measures such as Precision, recall, confusion matrix, ROC and F score. The experimental results illustrate that the proposed model outperforms in accuracy compared to existing approaches.

2. LITERATURE REVIEW

The section discusses the related work which has been carried out in the area of early prediction and diagnose of mental disorder.

The early prediction of mental illness of working professional is a novel research topic which is still under the investigation by many countries for several years. The problem of mental

illness will impact the working capacity of the professionals in a different way. It may lead to problems like absenteeism, reduction in productivity, disability[5].The long term illness may cause cardiovascular disease, skeleton disorders[6]. These mental disorders are mainly caused due to the fear of feeling and anxiety. The detailed literature review is carried out on anxiety disorder and a possible solution using 16 different machine learning algorithm. The paper came up with a conclusion that the existing technology may not be sufficient in predicting and diagnosing such disease and the novel idea is required for the better treatment of the disease[7]. A detailed study is carried out on mental illness and its possible impact. The study also listed possible ways to predict and control the disorder[8].The research is carried out to understand the possible relationship between mental illness and education, social-economic attainment[9].The investigation is made on the possible usage of machine learning algorithms such as SVM, K nearest neighbor in mental illness prediction[10].The hybrid combination of fuzzy logic with SVM is proposed combination for the diagnosis of mental disorder[11].The ontology-based multi labeled prediction model is proposed for the classification of therapeutic data[12].The study is made on the usage of a machine learning algorithm in suicide prevention[13].A detailed survey is carried out on the application of effective machine learning algorithm in the prevention of depressive disorder[14].The Bayesian network and SVM based machine learning algorithms are proposed for the prediction of mental illness. The SVM obtained good accuracy in comparison with Bayesian network algorithm[15].The artificial intelligent based framework is proposed to diagnose mental diseases. The algorithms such as Fuzzy logic, Fuzzy Genetic algorithm, and Rule-based systems are applied to treat the patients[16].A Neuro-Fuzzy based system is proposed for the prediction of mental depress among adults. The proposed system achieved an accuracy of 66%[17].The decision support based algorithm is proposed for the prediction of Schizophrenia disease[18].The backward strategy based depth-first search algorithm is proposed for the analysis of emotional symptoms and psychological diseases. The proposed method achieved an accuracy of 63%[19].The performance of the various classifier is proposed for the possible prediction of psychological disorder. The machine learning algorithms SVM, Random Forest, Expectation Maximum in combination with feature selection algorithms such as Relief and Information Gain has experimented. The experimental results illustrated that SVM in combination with information Gain obtained higher accuracy[20].The artificial Neural Network-based model is proposed to predict psychological disorder such as behavioral disorder, anxiety among the children. The proposed approach achieved an accuracy of 82%.The performance of various machine learning algorithm such as Artificial Network, Naïve Bayes classifier, and Decision Tree is analyzed in predicting Parkinson's disease. The Parkinson's is long term nervous system disorder which leads to anxiety, depression[21]. The novel framework is proposed using a probabilistic machine

learning algorithm for the diagnose the anxiety disorder. The supervised classification algorithms such as Naïve Bayes, Logistic Regression, and Bayesian Network are utilized in the early prediction of the disease[22].A Bayesian joint model is proposed for the prediction of anxiety disorder among adult[23].The experimental results illustrated that the proposed model achieved an AUC rate of 0.75 and ROC rate of 0.60.A Binary SVM based machine learning model was proposed for the prediction of general anxiety disorder among the youth[24]. The proposed system discriminated between healthy person and patient with depression. The system obtained an accuracy of 67%.Deep Learning-based predictive models are implemented for the prediction of autism. The ABIDE dataset was used to experiment. The proposed approach achieved an accuracy of 70%[25].There are few IoT based smart healthcare systems also developed in order to handle mental disorders[26].

3. PROPOSED METHODOLOGY

The paper proposes a novel Binary Logistic Gradient Transformation(BLGT) based Behavioral Model for the early prediction of mental health illness among working professionals. The dataset is accumulated from UCI repository containing details about 1260 working professional having 26 attributes in it. The prediction of health illness from the dataset is considered a binary classification problem. Hence the Binary response data will be optimally classified using BLGT model. The architecture diagram of the proposed BLGT model is shown in figure 1. Initially, the mental health data set is passed as input to the proposed system. The data set will be pre-processed and feature scaled to remove the noise outliers and missing values. The whole data set is scaled to normal range using the min-max scalar technique. The min-max normalization can be calculated using the equation (1) below.

$$f^1 = \frac{f - \min(f)}{\max(f) - \min(f)} \quad (1)$$

Where f is the original value of the given feature and f^1 is the normalized value of the feature. The mean of the normalization value is obtained using the equation (2) below.

$$f^1 = \frac{f - \text{average}(f)}{\max(f) - \min(f)} \quad (2)$$

The feature selection is considered an important stage in machine learning. The performance of the whole machine learning process depends on the techniques applied in the feature selection process. In the proposed approach the feature selection is performed using SDFS based algorithm. The dataset had a total of 26 features in it and after the process of feature selection, it has been reduced to 9. The logistic score of each and every feature is calculated to increase the maximum likely hood of the model at the same time to reduce the square error using the equation (3).

$$LS(Pb) = \delta + \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3 + \dots + \delta_n X_n \quad (3)$$

In equation (3) Pb is probability in the presence of features of interest. Instead of choosing the parameter which may raise

the square errors in the implemented model, the features are chosen which will increase the maximum likelihood of the model using the equation (4) below.

$$LS(Pb) = \ln\left(\frac{Pb}{1-Pb}\right) \tag{4}$$

Once the logistic scores are computed, model transformations are done by mapping the predicted values to the probabilities on the hypothesis space. The transformation will output the values which can be mapped between the binary classes using the sigmoid function. The steps involved in the model transformations are as below. The sigmoid function is represented using the equation (5) and (6).

$$f(k) = \frac{1}{1+e^{-k}} \tag{5}$$

The hypothesis can be derived from the sigmoid function using the formula below.

$$\sigma(K) = \sigma(\delta_0 + \delta_1 Y) \tag{6}$$

Hence the equation (7) can be represented as below.

$$K = \delta_0 + \delta_1 Y \tag{7}$$

Hence the probabilities are represented in hypothesis spaces '0' and '1' using the equation (8) and (9) below.

$$h\delta(y) = \text{sigmoid}(K) \tag{8}$$

$$h\delta(Y) = \frac{1}{(1+e^{-(\delta_0+\delta_1 Y)})} \tag{9}$$

Once the probability is represented in hypothesis space, using equation (10) and (11) model transformation results will map the results among binary classes using the following strategy.

$$Pb \geq 0.5, \text{class} : 1 \tag{10}$$

$$Pb < 0.5, \text{class} : 0 \tag{11}$$

The proposed approach imposes the penalty on the model based on the performance of the prediction using the Cross-Entropy function. The penalty is increased if the predicted probability differs from the actual label. In the proposed binary classification problem the Cross entropy will be imposed using the equation (12) below.

$$N = -(x \log(Pb) + (1-x) \log(1-Pb)) \tag{12}$$

In equation (12) above N indicates the number of classes (Presence of Mental illness, Absence of Mental Illness), the log is nothing but a natural log, x is a binary indicator(0 or 1) and Pb is Predicted probability by the model.

The cost of the penalty is minimized using Gradient descent function. It is one of the sophisticated first-order optimization algorithms which will minimize the cost by iteratively moving through the negative gradients. The gradient descent can be imposed using the equation below.

$$m^1(s) = m(s) (1 - m(s)) \tag{13}$$

$$Cw^1 = fv(m(s) - x) \tag{14}$$

In the equation (13) and (14) above m(s) is result of model prediction, fv is a feature vector, Cw is cost with respect to weights and x is an actual label with respect to class. The predicted probabilities are mapped to the corresponding

classes. The entire model is validated using 10 fold cross-validation technique to overcome the model overfitting and under the fitting issue. As shown in figure 1 , model is extensively trained using the mental health data set. The accuracy of the model is computed by passing the Test data set as input to the model. The accuracy obtained through the model is measured using the various parameter such as Precision, recall, confusion matric, ROC and F score. The process of training and testing the model is repeated until optimal accuracy is obtained. Given the new test, the BLGT model predicts the mental health illness of the Professional i.e, whether the professional need to undergo medical diagnosis or not. The Pseudocode of the proposed Binary Logistic Gradient Transformation Model is shown in the figure 2 .

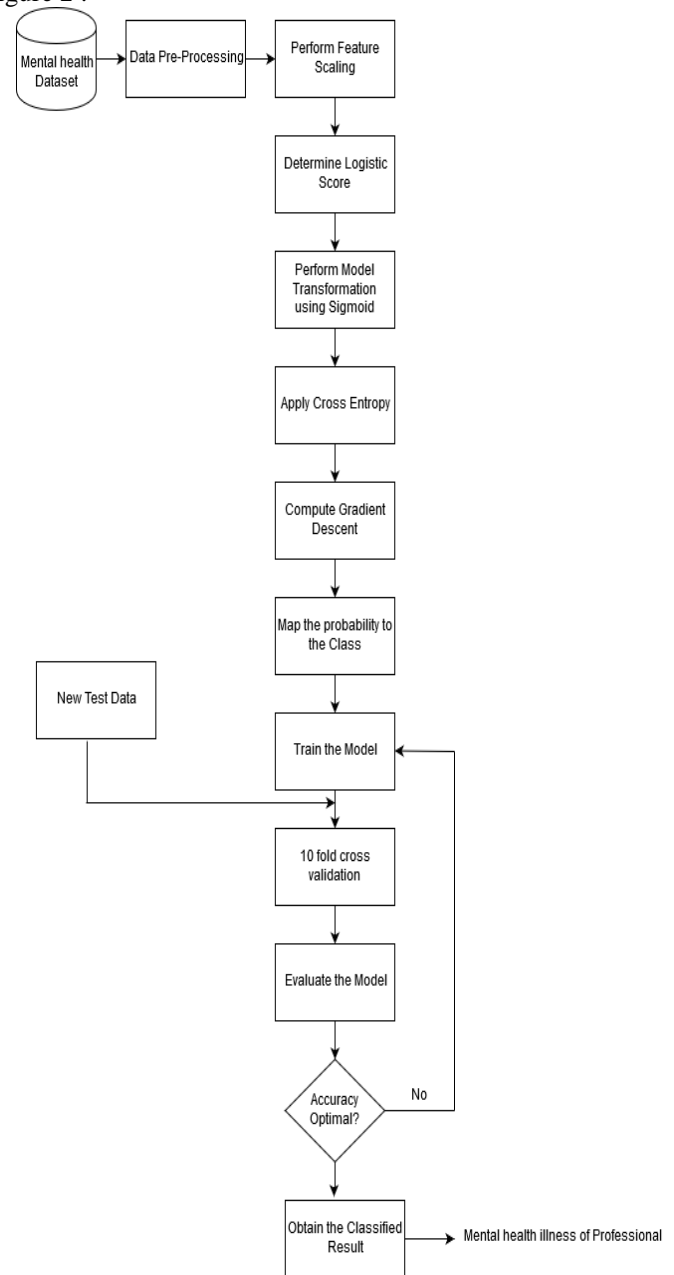


Figure 1: The architecture of the proposed BLGT approach

4. RESULT AND DISCUSSION

The experiment is conducted by passing the Mental health dataset as input to the model. The dataset is accumulated from the UCI repository contained details about 1260 working professional described. The Dataset had a total of 26 features in it. The few of the factors considered for predicting the mental health condition of professionalism. The input data is pre-processed to remove the noise and outlier. As a predictive model can understand only the numerical features, the whole dataset is encoded using Label encoder. The label encoder will convert all text into a model understandable format. The entire dataset is scaled and transformed into uniform range using min-max scalar. As part of data exploration, the probable cause of mental health disease with respect to different features age, Family history, Care options, Work interfere are analyzed as shown in figure 3 below.

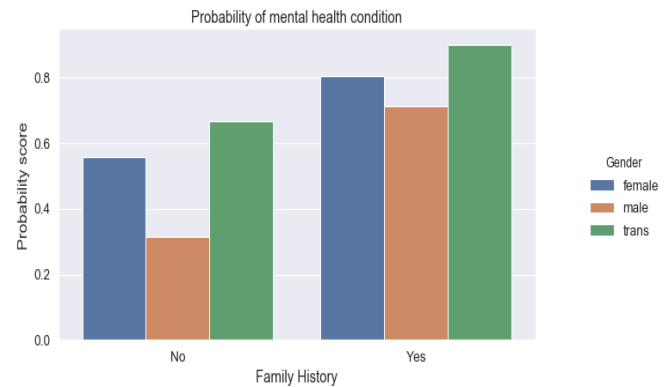


Figure 3: .Exploration analysis with respect to Family History

The SDFS based feature selection algorithm will select the top 8 features from the dataset which will have a direct impact on the prediction result. The proposed BLGT model is extensively trained on these selected features and performance of the model is analyzed by passing the Test data set as input to the proposed Model. The performance of the algorithm is validated using 10 fold cross-validation and measured using the various performance parameters such as Precision, Recall, F measure, AUC, accuracy and confusion matrix as shown in table 1.

Table 1 : Performance Parameter

Parameter	Formula
Recall(Rec)	$\frac{TP}{Actual\ Positive}$ $Actual\ Positive = (TP + FN)$
Accuracy	$\frac{TP + TN}{Total\ Population}$
F-score	$2 \times \frac{Prec \times Rec}{(Prec + Rec)}$
Precision(Prec)	$\frac{TP}{Actual\ Negative}$ $Actual\ Negative = (FP + TN)$

Pseudocode : Binary Logistic Gradient Transformation Model

Input :

Training Mental health dataset $F = (f_1, f_2, \dots, f_n)^T$

$F_{Test} = (f_1, f_2, \dots, f_n)$ //Feature values of predictor variables in test dataset

Output:

The Class of Test Dataset

Steps:

Let N be the number of classes

P(b) indicate the probability value

L(s) denote the Logistic score

m(s) denotes the result of model prediction ,

fv represent the feature vector,

Cw denotes cost with respect to weights

x is actual label with respect to class

1. Read the Training examples $F = (f_1, f_2, \dots, f_n)^T$ from Mental Health Dataset

2. Perform Feature scaling and Data Normalization on example attribute using

$$f^L = \frac{f - \min(f)}{\max(f) - \min(f)}$$

3. Determine the Logistic score of the probabilities using

$$LS(Pb) = \ln\left(\frac{Pb}{1-Pb}\right)$$

4. Use Sigmoid Function and Perform model Transformation using

$$f(k) = \frac{1}{1 + e^{-k}}$$

5. Calculate the Cross entropy

$$N = -(x \log(Pb) + (1 - x) \log(1 - Pb))$$

6. Determine the Gradient Descent using

$$m^L(s) = m(s) (1 - m(s))$$

$$Cw^L = fv(m(s) - x)$$

7. Represent the probability in hypothesis space and map the results of model transformation among binary classes results

8. Train the Logistic model using Data set $(f_1, f_2, \dots, f_n)^T$

9. Validate the model 10 fold cross validation

1. Divide the entire dataset in to 10 folds with equal approximate distribution of case and control

2. For each fold k_i out of K folds:

i. Set one of the fold k_i as test set

ii. Apply feature selection on remaining K-1 set

3. For the fold k_j in the K-1 folds:

i. Set the fold k_j as validation set

ii. Train the model with remaining set of K-2 fold

iii. Evaluate the performance of the model on k_j

4. Evaluate the performance of model on fold k_i

10. Repeat step 8 and 9 until accuracy is optimal

11. Get the Classified result on mental illness of

Figure 2: The Pseudocode of the proposed BLGT approach

To test the accuracy of the model, 378 instances containing the details of the working professionals are passed as input to the proposed model. Table 2 below indicates the details about the confusion matrix obtained with the proposed model. The table below indicates that out of 327 instances are classified correctly and 51 instances are classified incorrectly. Hence the proposed approach an accuracy of 87%.

Table 2: Confusion Matrix for the proposed BLGT Model

True Label	Negative	True Negative	153	False Positive	38
	Positive	False Negative	13	True Positive	174
		Negative		Positive	
		Predicted Label			

The AUC rate obtained with a proposed approach is shown in figure 4 below. The proposed approach achieved an AUC rate of 0.86 obtained with this model.

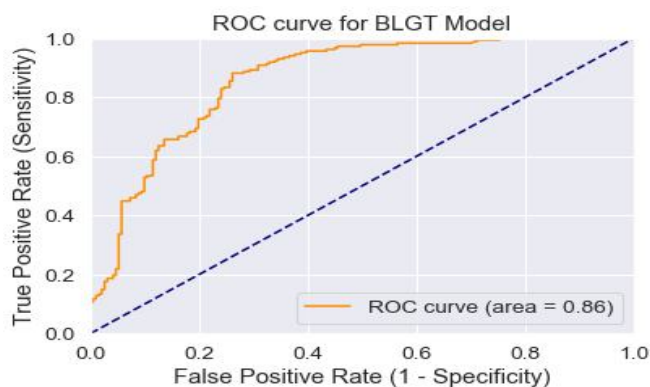


Figure 4 : ROC curve for proposed BLGT Model

The accuracy obtained with the proposed approach is compared with the base model such as SVM, Naïve Bayes, Decision Tree, and Logistic regression. The report of the comparative analysis is shown in Table 3 below

Table 3: Performance of Different Machine Learning Classifier

Technique	Precision	F measure	Accuracy	ROC	Cross validated AUC
Decision Tree	0.76	0.72	74	0.73	0.76
Random forest	0.82	0.79	81	0.79	0.75
SVM	0.74	0.74	83	0.81	0.84
Proposed BLGT model	0.85	0.89	87	0.86	0.87

Table 3 above demonstrates that the proposed BLGT model outperformed in performance compared to the existing approaches. The BLGT model achieved an outstanding cross-validated AUC rate of 0.87. This indicates that problems like model overfitting and underfitting are avoided in the proposed technique.

5. CONCLUSION

In this paper, a novel Binary Logistic Gradient Transformation(BLGT) based Behavioral Model has been proposed for the Early Mental Health Illness Diagnosis among working Professionals. The Mental health dataset is pre-processed and important features which will impact the prediction decision are selected using the SDFS based feature selection algorithm. The Logistic model is implemented and transformed using the sigmoid function. The penalty is imposed on the model using Cross entropy loss function. The model is optimized using a Gradient Descent function. The model is validated using 10 fold cross-validation technique to overcome the model overfitting issue. The accuracy obtained through the proposed BLGT model is compared using various machine learning classifier. The experimental results demonstrate that the proposed BLGT model is outperformed in accuracy compared to existing approaches.

REFERENCES

1. Chekroud, Sammi R., et al. "Association between physical exercise and mental health in 1· 2 million individuals in the USA between 2011 and 2015: a cross-sectional study." *The Lancet Psychiatry* 5.9 (2018): 739-746.
2. Hegde, Sandeepkumar, and Monica R. Mundada, "Symmetric disparity based Feature selection Approach for Chronic Disease Data Analysis ", 2019 6th Indian Symposium on Computer Systems, [HPC-India] Indian Institute of Science, Bangalore, July 19-20.
3. Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: the University of California, School of Information and Computer Science.
4. Yin, Xinyou, et al. "A flexible sigmoid function of determinate growth." *Annals of botany* 91.3 (2003): 361-371.
5. Banerjee, Chaity, Tathagata Mukherjee, and Eduardo Pasillao Jr. "An Empirical Study on Generalizations of the ReLU Activation Function." *Proceedings of the 2019 ACM Southeast Conference*. ACM, 2019. <https://doi.org/10.1145/3299815.3314450>
6. Kazdin, Alan E. "Annual Research Review: Expanding mental health services through novel models of intervention delivery." *Journal of Child Psychology and Psychiatry* 60.4 (2019): 455-47
7. Spielberg, Charles D. "8 Effects of stress, emotion, and Type-A behavior on heart disease and

- psychological disorders.**" Progress in Psychological Science Around the World. Volume 2: Social and Applied Issues: Proceedings of the 28th International Congress of Psychology. Routledge, 2019.
8. Chapman, Benjamin P., et al. "**Health risk prediction models incorporating personality data: Motivation, challenges, and illustration.**" *Personality Disorders: Theory, Research, and Treatment* 10.1 (2019): 46.
 9. Van Bronswijk, Suzanne C., et al. "**A prognostic index for long-term outcome after successful acute phase cognitive therapy and interpersonal psychotherapy for major depressive disorder.**" *Depression and anxiety* 36.3 (2019): 252-261.
<https://doi.org/10.1002/da.22868>
 10. Anggoro, Dimas Aryo, and Naqshauliza Devi Kurnia. "**Comparison of Accuracy Level of Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) Algorithms in Predicting Heart Disease.**" *International Journal of Emerging Trends in Engineering Research*, Volume 8, no. 5 (2020).
<https://doi.org/10.30534/ijeter/2020/32852020>
 11. Srividya, M., S. Mohanavalli, and N. Bhalaji. "**Behavioral modeling for mental health using machine learning algorithms.**" *Journal of medical systems* 42.5 (2018): 88.
 12. Srividya, M., S. Mohanavalli, and N. Bhalaji. "**Behavioral modeling for mental health using machine learning algorithms.**" *Journal of medical systems* 42.5 (2018): 88
 13. Wang, Haoxiang, and Jingbin Wang. "**An effective image representation method using kernel classification.**" 2014 IEEE 26th international conference on tools with artificial intelligence. IEEE, 2014.
<https://doi.org/10.1109/ICTAI.2014.131>
 14. Ribeiro, J. D., et al. "**Letter to the Editor: suicide as a complex classification problem: machine learning and related techniques can advance suicide prediction-a reply to Roaldset (2016).**" *Psychological medicine* 46.9 (2016): 2009-2010.
 15. Kessler, Ronald C., et al. "**Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports.**" *Molecular psychiatry* 21.10 (2016): 1366.
 16. Bakar, Wan Aezwani Wan Abu, Mustafa Man, Wan Suryani Wan Awang, Nur Laila NajwaMohd Josdi, AimanNatasya Naquiyah, and Nik Nurul Nadia Nik Pa. "**HDP: Heart Disease Prediction Tool using Neural Network.**" *International Journal of Emerging Trends in Engineering Research*, Volume 8, no. 5 (2020).
<https://doi.org/10.30534/ijeter/2020/50852020>
 17. Luxton, David D. "**Artificial intelligence in psychological practice: Current and future applications and implications.**" *Professional Psychology: Research and Practice* 45.5 (2014): 332.
 18. Chattopadhyay, Subhagata, et al. "**An automated system to diagnose the severity of adult depression.**" 2011 Second International Conference on Emerging Applications of Information Technology. IEEE, 2011.
 19. Comin Nunes, Luciano, et al. "**Handling diagnosis of schizophrenia by a hybrid method.**" *Computational and mathematical methods in medicine* 2015 (2015).
 20. Nagar, T. "**Prediction of Mental Health Problems Among Children Using Machine Learning Techniques.**" *NIMHANS Journal.*, 14(2): 99-106, 1996 Apr
 21. Kipli, Kuryati, Abbas Z. Kouzani, and Isredza Rahmi A. Hamid. "**Investigating machine learning techniques for detection of depression using structural MRI volumetric features.**" *International journal of bioscience, biochemistry and bioinformatics* 3.5 (2013): 444-448.
 22. Dabek, Filip, and Jesus J. Caban. "**A neural network based model for predicting psychological conditions.**" *International conference on brain informatics and health.* Springer, Cham, 2015.
 23. Chen, Henian, Yangxin Huang, and Nanhua Zhang. "**Joint modeling of a linear mixed effects model for self-esteem from mean ages 13 to 22 and a generalized linear model for anxiety disorder at mean age 33.**" *Journal of Medical Statistics and Informatics* 3.1 (2015).
<https://doi.org/10.7243/2053-7662-3-1>
 24. Chatterjee, Moitreyee, et al. "**Context-based signal descriptors of heart-rate variability for anxiety assessment.**" 2014 IEEE international conference on acoustics, speech and signal processing (icassp). IEEE, 2014.
 25. Hilbert, Kevin, et al. "**Separating generalized anxiety disorder from major depression using clinical, hormonal, and structural MRI data: A multimodal machine learning study.**" *Brain and Behavior* 7.3 (2017): e00633.
 26. Reddy, K. Uday Kumar, S. Shabbiha, and M. Rudra Kumar. "**Design of High Security Smart Health Care Monitoring System using IoT.**" *International Journal of Emerging Trends in Engineering Research*, Volume 8, no. 6 (2020).
<https://doi.org/10.30534/ijeter/2020/09862020>