International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

Measuring Mental Health Condition using Logistic regression

Anubrata Deb[1], Bristi Samadder[2], Souroja Chowdhury[3], Mrs. Sulekha Das[4], Mrs. Shweta Banarjee[5]
UG-Student, Computer Science and Engineering, Techno Engineering College Banipur[1]
UG-Student, Computer Science and Engineering, Techno Engineering College Banipur[2]
UG-Student, Computer Science and Engineering, Techno Engineering College Banipur[3]
Assistant Professor, Techno Engineering College Banipur[4]
Assistant Professor, Techno Engineering College Banipur[5]
***CORRESPONDING AUTHOR ORCID ID:** 0000-0003-4138-7400 [1], 0009-0009-2660-6899 [2], 0009-0005-3443-3968 [3], 0000-0002-6641-3268 [4], 0000-0002-7888-3136 [5]*

Abstract: Psychiatry begun to develop empirical approaches only in the past 30 years to conceptualizing, assessing and documenting positive mental health. In society we accept mental disorders as normal and try to hide it. The author has come across the fact that how severe Mental illness affect our life and the measures that can be taken to heal the same. Authors have designed means of prediction of depression of an individual. Real data has been collected from people ranging from 18 to 59 years of age. Neurotic sufferings cannot be fully understood without understanding real health and sound approach towards life so, the authors have taken approach of all the factors including depression, anxiety, stress, eating disorder, short temper, hallucination, loneliness, suicidal attempt, constant guilt feeling, fearful all these and more, realistic disturbing state and trait that result in various mental diseases has been considered and analysed. Authors have used logistic regression model on the study they have come across for predicting depression of an individual.
Keywords: Logistic regression, Mental Health, Prediction of Depression, Classification, Dataset, Machine Learning.

1.INTRODUCTION

In our society most of the people suffers from depression. Depressed person can be a family person or an employee and the severe effect can be seen in every aspect of life due to depression. Authors have taken 40-50 attributes on which it has been observed that one goes into depression including one or multiple attributes. Authors have also observed that many of the depressed people have gone to psychiatry and have healed them completely but at the same time the new age generation have taken the help of our tradition like Yoga and meditation and have got freedom from depression. However, some people face depression throughout their life and no healing approach can rectify the same. So, it is oblivious that depression is becoming very challenging mental health issue and the authors have found all these attributes including childhood abuse, heart break, divorce, past trauma and many more realistic challenges lies behind the depression.
Nowadays, a statistical analysis is widely used in various fields such as in science, medicine and in social sciences too. Regression is one of the most common statistical methods used. Logistic regression (LR) is such a statistical technique that helps to determine a mathematical relation between the multiple independent variables and a single dependent variable. We aim to develop prediction of depression that would predict the probability of a person getting depressed based on their mental state and response that they generate as a result of those attributes like emotional over eating, sleeplessness and many more. The prediction of depression is done by machine learning using stepwise logistic regression analysis. Based on the analysis, it predicts the depression of every individual.

2.LITERACY REVIEW

Numerous studies on automatic detection of the symptoms of depression have been carried out using ‘Machine learning’ methods such as ‘Supervised learning’ such as ‘Logistic regressions’. Logistic regressions are an applied math technique, the target of Logistic regression analysis is to use the

Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

independent variables whose worth square measure acknowledged predicting the worth of the only dependent value. Data has been collected through Google form by circulating this. In researchers we have also focused on depression detection using text messages from social media platforms, such as Twitter, WhatsApp -in the hope that social media texts can help detect depression even when the individual is unaware of their depression or is in denial. The majority of research studies on depression detection using social media messages usually follow either a textual-based featuring approach or a person descriptive-based featuring approach The main aim of this project is the prediction of the dependence of depression in the society on Age, Gender, Occupation, Relation status, Health issues, Family history and Past incidents. These features are then input into the detection models. Most models for depression detection have been developed using multiple regressions classifiers, such as the Support Confusion matrixes such as True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Sensitivity, Specificity, Precision, Recall and Accuracy.
These are the research papers that I have taken help form

Burdisso et al. (2019) [9]
Burdisso et al. used logistic regression to predict depression in a sample of adolescents based on features such as the use of negative or emotional language. The study found that the logistic regression model had an accuracy of 82.8% in detecting depression.

Jung et al. (2017) [10]
Jung et al. developed a logistic regression model to predict depression from social media data which is called sense mood. The approach, which involves the creation of a vocabulary of relevant concepts and relationships, to analyse the sentiment and context of social media posts. The study found that the logistic regression model had an accuracy of 75% in detecting depression.

Lin et al. (2020) [12]
Lin et al. developed a logistic regression model to predict depression from social media data which is called sense mood. The approach uses various features extracted from social media posts such as linguistic and emotional features, as well as user engagement and network features. The study found that the logistic regression model had an accuracy of 81.2% in detecting depression.

Chen et al. (2018) [7]
Chen et al. developed a logistic regression model to predict depression in a sample of patients with cardiovascular disease. The model used demographic variables, clinical variables, and self-reported symptoms as predictors. The study found that the logistic regression model had an accuracy of 75.1% in detecting depression.

Table.1. *An overview of related works:*

Dataset	Source	Total records	Depression records		Non-Depression records	
			Total	%	Total	%
Shen et al. [2]	Twitter	11877	6493	54.67	5384	45.33
Eye [1]	Twitter	10314	2314	22.44	8000	77.56
Tanwar [4]	Victoria’s diary	62	62	100	0	0
Virahonda [5]	Facebook	9178	9178	100	0	0
Komati [6]	Reddit about depression	50000	50000	100	0	0

3.METHODOLOGY
In this paper, data were taken from our own network such as relative, friends circle and using social media.

DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

Table.2. *Dataset description:*

No	ATTRIUTES	DESCRIPTION
1.	Age	Age of the person
2.	Gender	Gender of the person
3.	Occupation	Occupation of the person
4.	Relation	Relation status of the person
5.	Insecurity	Mainly faced problems by the person
6.	Irritability	Mainly faced problems by the person
7.	Loneliness	Mainly faced problems by the person
8.	Convolution	Mainly faced problems by the person
9.	Hallucination	Mainly faced problems by the person
10.	Thought of self-harm or suicide	Mainly faced problems by the person
11.	Aggression	Mainly faced problems by the person
12.	Nervousness or excessive sweating	Mainly faced problems by the person
13.	Lack of concentration or confidence	Mainly faced problems by the person
14.	Anything doing more than it's needed	Mainly faced problems by the person
15.	Difficulty in breathing or shortness in breath	Mainly faced problems by the person
16.	Fatigue	Other problems faced by the person
17.	Insomnia	Other problems faced by the person
18.	Restlessness	Other problems faced by the person
19.	Loss of appetite	Other problems faced by the person
20.	Intrusive thoughts	Other problems faced by the person
21.	Guilt without any reason	Other problems faced by the person
22.	Sudden weights gain or loss	Other problems faced by the person
23.	Excessive crying or eating or sleeping	Other problems faced by the person
24.	Hopelessness or loss of interest in activities	Other problems faced by the person
25.	Drug and alcohol missuses	Reasons of these problems
26.	Excessive family pressure	Reasons of these problems
27.	Severe or long-term stress	Reasons of these problems
28.	Social isolation or loneliness	Reasons of these problems
29.	Homelessness or poor housing	Reasons of these problems
30.	Unemployment or losing your job	Reasons of these problems
31.	Being the victim of a violent crime	Reasons of these problems
32.	Excessive mental stress for studies	Reasons of these problems
33.	Childhood abuse, trauma or neglect	Reasons of these problems
34.	Social disadvantage, poverty or debt	Reasons of these problems
35.	Being a long-term carrier for someone	Reasons of these problems
36.	Insecurity for excessive self-obsession	Reasons of these problems
37.	Bereavement (loosing someone close to you)	Reasons of these problems
38.	Having a long-term physical health condition	Reasons of these problems
39.	Domestic violence, bullying or other abuses as an adult	Reasons of these problems
40.	Experiencing discrimination and stigma, including racism	Reasons of these problems
41.	Being involve in a serious incident in which you feared for your life	Reasons of these problems
42.	Physical causes for example a head injury or a neurological condition	Reasons of these problems
43.	Family history	Family history of the person

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

Multiple linear regression is a statistical technique used to model the relationship between a dependent variable and multiple independent variables. It assumes a linear relationship between the dependent variables and each independent variable, and seeks to find the coefficients that best describe this relationship.

Logistic regression is a statistical method used to analyse the relationship between a categorical dependent variable and one or more independent variables. It is a type of regression analysis that is used when the dependent variable is binary or dichotomous (i.e., it can take only two possible values). In logistic regression, the response variable is usually coded as 1 for success or 0 for failure, and the predictor variables can be continuous or categorical. The logistic regression model uses a mathematical function called the logistic function, which transforms the linear combination of predictor variables into a value between 0 and 1 that represents the probability of success. Here we detect the depression is occur or not (Yes=1, No=0).

While multiple linear regression seeks to predict a continuous outcome variable, logistic regression is used to predict the probability of an event occurring. Therefore, logistic regression is commonly used for binary classification problems, such as predicting whether a customer will buy a product or not, while multiple linear regression can be used for predicting continuous outcomes such as predicting the house prices.

3.1 Accuracy of difference between Actual data and Calculated data

In this research, the hypotheses that were used:
H0: b1=b2=b3=b4=0
Ha: At least one of the b1, b2, b3, and b4 does not equal 0 which says that
H0: None of the controlled variables X1, X2, X3, and X4 is significantly related to Y
Ha: At least one of the controlled variables X1, X2, X3, and X4 is significantly related to Y The model of multiple regression can be represented as:

$$\boldsymbol{Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + \dots \dots \dots \dots + b_nX_n}$$

Here,
Y= Dependent Variable (Depression),
a= Constant variable,
b_1= Coefficient of the first control variable,
b_2= Coefficient of the second control variable,
b_3= Coefficient of the third control variable,
b_4= Coefficient of the fourth control variable,
b_5= Coefficient of the fifth control variable,
b_6= Coefficient of the sixth control variable,
b_7= Coefficient of the seventh control variable,
b_8= Coefficient of the eighth control variable,
b_9= Coefficient of the ninth control variable,
b_{10}= Coefficient of the tenth control variable,
b_{11}= Coefficient of the eleventh control variable,
b_{12}= Coefficient of the twelfth control variable,
b_{13}= Coefficient of the thirteenth control variable,
b_{14}= Coefficient of the fourteenth control variable,
b_{15}= Coefficient of the fifteenth control variable,
b_{16}= Coefficient of the sixteenth control variable,
b_{17}= Coefficient of the seventeenth control variable,
b_{18}= Coefficient of the eighteenth control variable,
b_{19}= Coefficient of the nineteenth control variable,
b_{20}= Coefficient of the twentieth control variable,

DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

b_{21}= Coefficient of the twenty-first control variable,
b_{22}= Coefficient of the twenty-second control variable,
b_{23}= Coefficient of the twenty-third control variable,
b_{24}= Coefficient of the twenty-fourth control variable,
b_{25}= Coefficient of the twenty-fifth control variable,
b_{26}= Coefficient of the twenty-sixth control variable,
b_{27}= Coefficient of the twenty-seventh control variable,
b_{28}= Coefficient of the twenty-eighth control variable,
b_{29}= Coefficient of the twenty-ninth control variable,
b_{30}= Coefficient of the thirtieth control variable,
b_{31}= Coefficient of the thirty-first control variable,
b_{32}= Coefficient of the thirty-second control variable,
b_{33}= Coefficient of the thirty-third control variable,
b_{34}= Coefficient of the thirty-fourth control variable,
b_{35}= Coefficient of the thirty-fifth control variable,
b_{36}= Coefficient of the thirty-sixth control variable,
b_{37}= Coefficient of the thirty-seventh control variable,
b_{38}= Coefficient of the thirty-eighth control variable,
b_{39}= Coefficient of the thirty-ninth control variable,
b_{40}= Coefficient of the fortieth control variable,
x_1= Controlled variable (Age),
x_2= Controlled variable (Gender),
x_3= Controlled variable (Occupation),
x_4= Controlled variable (Relation),
x_5= Controlled variable (Insecurity),
x_6= Controlled variable (Irritability),
x_7= Controlled variable (Loneliness),
x_8= Controlled variable (Convolution),
x_9= Controlled variable (Hallucination),
x_{10}= Controlled variable (Thought of self-harm or suicide),
x_{11}= Controlled variable (Aggression),
x_{12}= Controlled variable (Nervousness or excessive sweating),
x_{13}= Controlled variable (Lack of concentration or confidence),
x_{14}= Controlled variable (Anything doing more than it's needed),
x_{15}= Controlled variable (Difficulty in breathing or shortness in breath),
x_{16}= Controlled variable (Fatigue),
x_{17}= Controlled variable (Insomnia),
x_{18}= Controlled variable (Restlessness),
x_{19}= Controlled variable (Loss of appetite),
x_{20}= Controlled variable (Intrusive thoughts),
x_{21}= Controlled variable (Guilt without any reason),
x_{22}= Controlled variable (Sudden weights gain or loss),
x_{23}= Controlled variable (Excessive crying or eating or sleeping),
x_{24}= Controlled variable (Hopelessness or loss of interest in activities),
x_{25}= Controlled variable (Drug and alcohol missuses),
x_{26}= Controlled variable (Excessive family pressure),
x_{27}= Controlled variable (Severe or long-term stress),
x_{28}= Controlled variable (Social isolation or loneliness),
x_{29}= Controlled variable (Homelessness or poor housing),
x_{30}= Controlled variable (Unemployment or losing your job),
x_{31}= Controlled variable (Being the victim of a violent crime),

DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

x_{32}= Controlled variable (Excessive mental stress for studies),
x_{33}= Controlled variable (Childhood abuse, trauma or neglect),
x_{34}= Controlled variable (Social disadvantage, poverty or debt),
x_{35}= Controlled variable (Being a long-term carrier for someone),
x_{36}= Controlled variable (Insecurity for excessive self-obsession),
x_{37}= Controlled variable (Bereavement (loosing someone close to you)),
x_{38}= Controlled variable (Having a long-term physical health condition),
x_{39}= Controlled variable (Domestic violence, bullying or other abuses as an adult),
x_{40}= Controlled variable (Experiencing discrimination and stigma, including racism),
x_{41}= Controlled variable (Being involve in a serious incident in which you feared for your life),
x_{42}= Controlled variable (Physical causes for example a head injury or a neurological condition),
x_{42}= Controlled variable (Family history)

H0: The logistic regression is presented as:

$\mathbf{Y_1 = Y/(1+e^{-Y})}$
Here,
Y= Dependent Variable
e = Euler's number

3.2 Confusion-Matrix

After finding the accuracy of the difference between actual data and calculated data we did the Confusion Matrix. In this confusion matrix it can be seen that, we find the TP – which stands for 'TRUE POSITIVE' means the accuracy of classified positive data, TN – which stands for 'TRUE NEGATIVE' means the accuracy of classified negative data, FP – which stands for 'FALSE POSITIVE', means which remark that actual value is negative but predicted data is positive, FN – which stands for 'FALSE NEGATIVE' means that actual data and the predicted data both are negative and append the TP, TN, FP, FN value in 2*2 matrix. After that, we find the accuracy, sensitivity, precision, recall, and specificity. This matrix contains all the raw information about the predictions done by a classification model on a given data set.

3.3 Cross-Validation

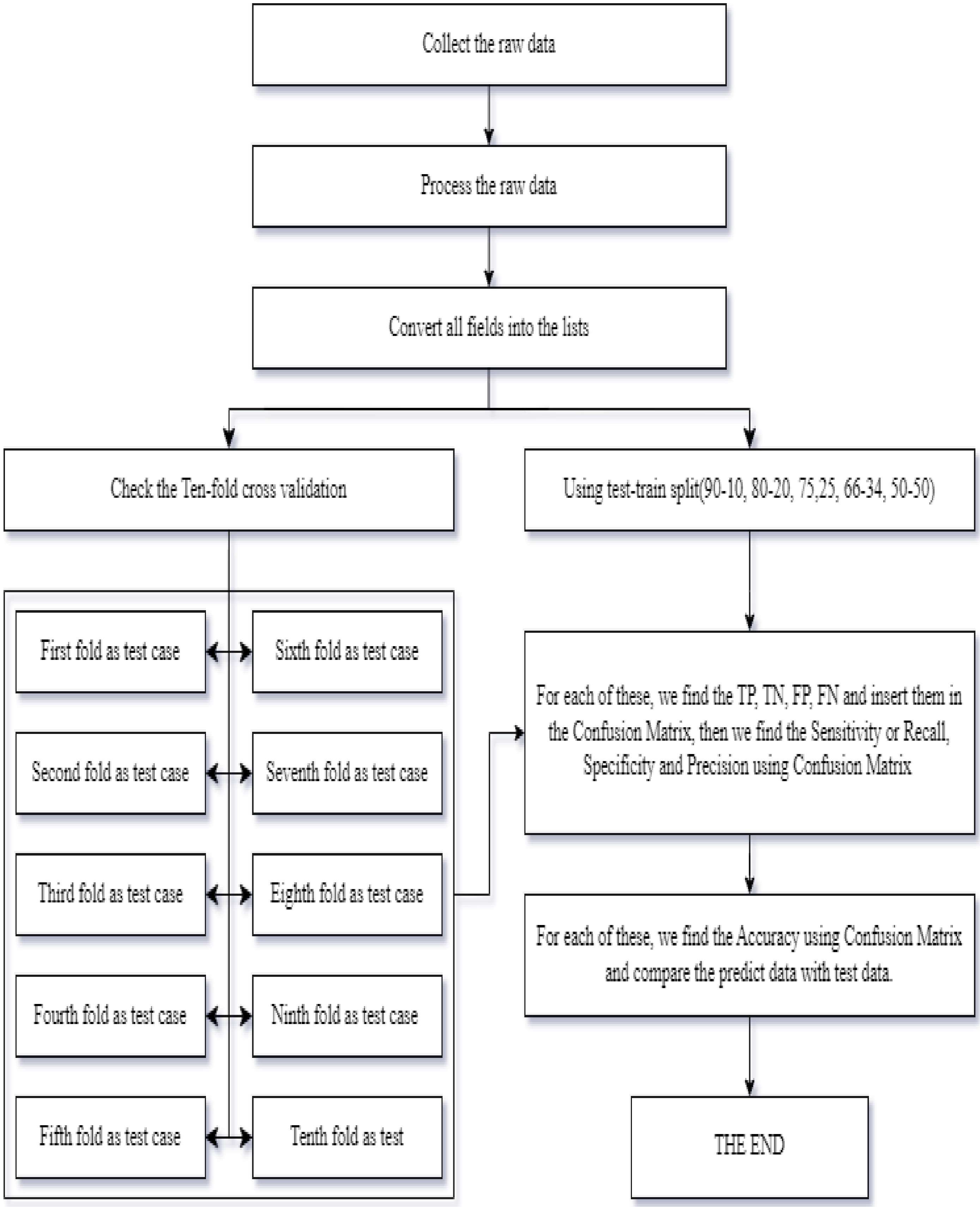
After finding the accuracy of the difference between actual data and calculated data we did cross-validation. In this cross-validation process first, we divide the whole list into 10 sub-list and then we find the accuracy of 10 sub-list elements we also find the Confusion Matrix of each Sub-list and we find the accuracy, and sensitivity, precision, recall, and specificity.

❑ **ACCURACY:** It's the ratio of the correctly labelled subjects to the whole pool of subjects. Accuracy is intuitional.
❑ **PRECISION:** Precision is the ratio of the correctly +ve labelled by our program to all +ve labelled.
❑ **RECALL:** Recall means out of the total positive, what percentage are predicted positive.
❑ **SPECIFICITY:** Specificity is calculated as the number of correct negative predictions divided by the total number of negatives.

- ACCURACY= (TP+TN/ TP+TN+FP+FN) * 100
- PRECISION = (TP/FP+TP) *100
- RECALL= (TP/FN+TP) *100
- SPECIFICITY = (TN/TN+FP) * 100

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

3.4 Flowchart



DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

4.RESULT AND DISCUSSION

Table.3. *Accuracy of difference between Actual data and Calculated data:*

Accuracy of 90%Data as Training Data or (0.90)	90.48
Accuracy of 80%Data as Training Data or (0.80)	73.17
Accuracy of 75%Data as Training Data or (0.75)	67.31
Accuracy of 66%Data as Training Data or (0.66)	61.97
Accuracy of 50%Data as Training Data or (0.50)	61.32

Table.4. *Confusion Matrix & Corresponding Result:*

For 90% of Data

Confusion Matrix:

18	2
0	1

Accuracy:	90.48
Precision:	90.00
Recall:	100.0
Specificity:	33.33

For 80% of Data

Confusion Matrix:

25	11
0	5

Accuracy:	73.17
Precision:	69.44
Recall:	100.0
Specificity:	31.25

For 75% of Data

Confusion Matrix:

30	17
0	5

Accuracy:	67.31
Precision:	63.83
Recall:	100.0
Specificity:	22.73

For 66% of Data

Confusion Matrix:

43	27
0	1

Accuracy:	61.97
Precision:	61.43
Recall:	100.0
Specificity:	03.57

For 50% of Data

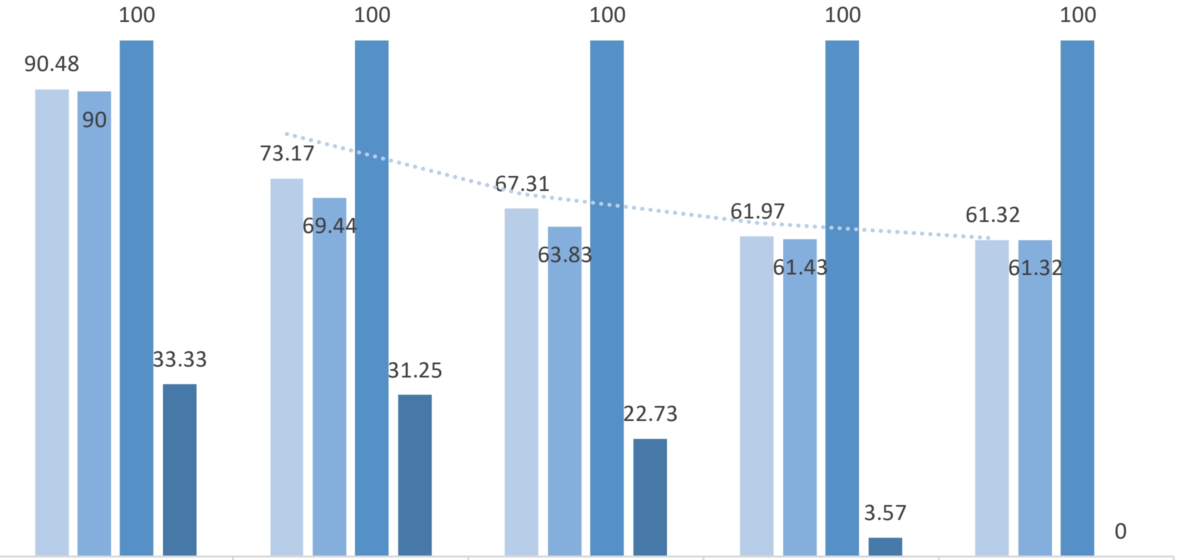
Confusion Matrix:

65	41
0	0

Accuracy:	61.32
Precision:	61.32
Recall:	100.0
Specificity:	00.00

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

Graph.1. *Confusion matrix graph for varies test-train split data set:*



	90% TRAINING DATA SET	80% TRAINING DATA SET	75% TRAINING DATA SET	66% TRAINING DATA SET	50% TRAINING DATA SET
Accuracy	90.48	73.17	67.31	61.97	61.32
Precision	90	69.44	63.83	61.43	61.32
Recall	100	100	100	100	100
Specificity	33.33	31.25	22.73	3.57	0

Table.5. *For Ten-fold cross-validation Accuracy:*

TEST CASE	ACCURACY RATE (%)
1	75.00
2	65.00
3	85.00
4	85.00
5	75.00
6	75.00
7	60.00
8	55.00
9	60.00

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

10	86.96

First Fold

Confusion Matrix:

13	5
0	2

Accuracy:	75.00
Precision:	72.22
Recall:	100.0
Specificity:	28.57

Second Fold

Confusion Matrix:

8	7
0	5

Accuracy:	65.00
Precision:	53.33
Recall:	100.0
Specificity:	41.67

.

Third Fold

Confusion Matrix:

10	3
0	7

Accuracy:	85.00
Precision:	76.92
Recall:	100.0
Specificity:	70.00

Fourth Fold

Confusion Matrix:

17	3
0	0

Accuracy:	85.00
Precision:	85.00
Recall:	100.0
Specificity:	00.00

Fifth Fold

Confusion Matrix:

15	5
0	0

Accuracy:	75.00
Precision:	75.00
Recall:	100.0
Specificity:	00.00

Sixth Fold

Confusion Matrix:

15	5
0	0

Accuracy:	75.00
Precision:	75.00
Recall:	100.0
Specificity:	00.00

Seventh Fold

Confusion Matrix:

12	8
0	0

Accuracy:	60.00
Precision:	60.00
Recall:	100.0
Specificity:	00.00

Eighth Fold

Confusion Matrix:

10	9
0	1

Accuracy:	55.00
Precision:	52.63
Recall:	100.0
Specificity:	10.00

Table.6. *For Ten-fold cross-validation Results:*

Ninth Fold

Confusion Matrix:

8	8
0	4

Accuracy:	60.00
Precision:	50.00
Recall:	100.0
Specificity:	33.33

Tenth Fold

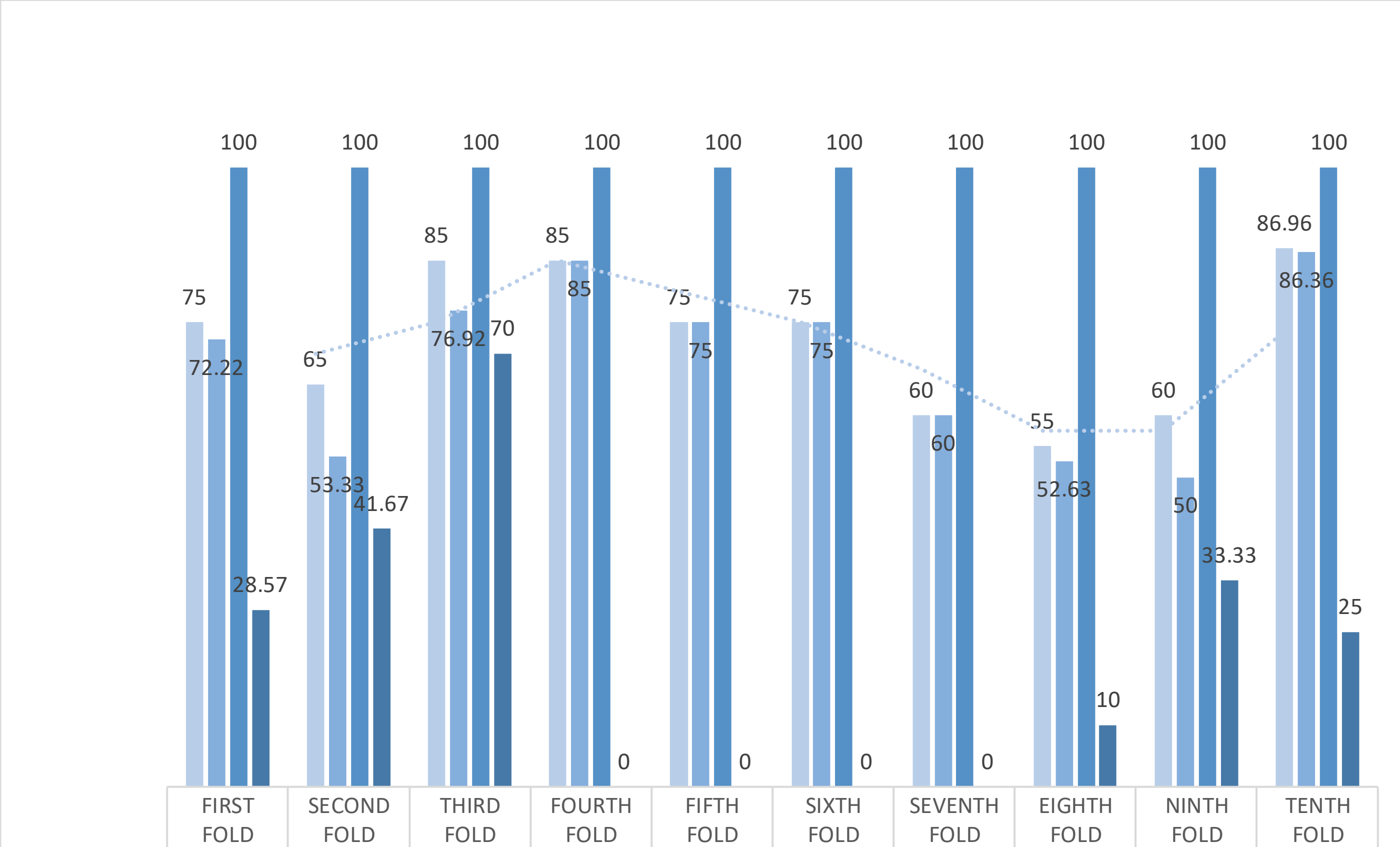
Confusion Matrix:

19	3
0	1

Accuracy:	86.96
Precision:	86.36
Recall:	100.0
Specificity:	25.00

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

Graph.2. *Confusion matrix graph for Ten-fold cross-validation:*



	FIRST FOLD	SECOND FOLD	THIRD FOLD	FOURTH FOLD	FIFTH FOLD	SIXTH FOLD	SEVENTH FOLD	EIGHTH FOLD	NINTH FOLD	TENTH FOLD
Accuracy	75	65	85	85	75	75	60	55	60	86.96
Precision	72.22	53.33	76.92	85	75	75	60	52.63	50	86.36
Recall	100	100	100	100	100	100	100	100	100	100
Specificity	28.57	41.67	70	0	0	0	0	10	33.33	25

5.CONCLUSIONS

This paper uses Logistic regressions to predict the depression. It has proved to be a powerful statistical tool for measuring mental health conditions. Through the use of this method, researchers can identify the risk factors and predictors of various mental health conditions, which can lead to the development of more effective prevention and treatment strategies. Moreover, logistic regression can help in the early detection of mental health conditions, which is crucial in improving health outcomes and preventing long-term disability. We have collected the data by circulating the google form among our relatives, collages, teachers and friends based on that we made a relationship between the dependent variable and the independent variable after that we perform Confusion Matrix where we compare the actual target values with those predicted by the machine learning model. After checking the Confusion Matrix, we move to the Cross Validation where we find the accuracy of 10 sub-list elements and we also find the Confusion Matrix of each Sub-list. we predict the accuracy as well as sensitivity, precision, recall, and specificity for user choice test data and the 10 sub-list.

The results of this study suggest that logistic regression can be used as a reliable tool for measuring mental health conditions. However, it is important to note that the accuracy of the results is highly dependent on the quality of data collected, the sample size, and the choice of variables. Thus, further

DOI:10.46647/ijetms.2023.v07i02.040 ISSN: 2581-4621

research is needed to enhance the predictive power of logistic regression models for mental health conditions.

6.REFERENCES

1. Chiong, R., Budhi, G. S., Dhakal, S., & Chiong, F. (2021). A textual-based featuring approach for depression detection using machine learning classifiers and social media texts. Computers in Biology and Medicine, 135, 104499.
2. Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., ... & Zhu, W. (2017, August). Depression detection via harvesting social media: A multimodal dictionary learning solution. In IJCAI (pp. 3838-3844).
3. Tanwar, R. (2020). Victoria Suicide Data.
4. Virahonda, S. (2020). Depression and anxiety comments.
5. Komati, N. (2020). r/SuicideWatch and r/depression posts from Reddit.
6. Chen, Y., Zhou, B., Zhang, W., Gong, W., & Sun, G. (2018, June). Sentiment analysis based on deep learning and its application in screening for perinatal depression. In 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC) (pp. 451-456). IEEE.
7. Islam, M. R., Kabir, M. A., Ahmed, A., Kamal, A. R. M., Wang, H., & Ulhaq, A. (2018). Depression detection from social network data using machine learning techniques. Health information science and systems, 6, 1-12.
8. Burdisso, S. G., Errecalde, M., & Montes-y-Gómez, M. (2019). A text classification framework for simple and effective early depression detection over social media streams. Expert Systems with Applications, 133, 182-197.
9. Farima, I. (2019). Ud Din Abbasi B, Khan S, Al-Saeed M, Farroq Ahmad H, Mumtaz R. Prediction of postpartum depression using machine learning techniques from social media text. Expert Systems, 36(4).
10. Jung, H., Park, H. A., & Song, T. M. (2017). Ontology-based approach to social data sentiment analysis: detection of adolescent depression signals. Journal of medical internet research, 19(7), e259.
11. Lin, C., Hu, P., Su, H., Li, S., Mei, J., Zhou, J., & Leung, H. (2020, June). Sensemood: depression detection on social media. In Proceedings of the 2020 international conference on multimedia retrieval (pp. 407-411).
12. Daly, M., & Robinson, E. (2021). Psychological distress and adaptation to the COVID-19 crisis in the United States. Journal of psychiatric research, 136, 603-609.
13. Laali, K. K., Greves, W. J., Correa-Smits, S. J., Zwarycz, A. T., Bunge, S. D., Borosky, G. L., ... & Chanan-Khan, A. (2018). Novel fluorinated curcuminoids and their pyrazole and isoxazole derivatives: Synthesis, structural studies, Computational/Docking and in-vitro bioassay. Journal of Fluorine Chemistry, 206, 82-98.
14. Priya, A., Garg, S., & Tigga, N. P. (2020). Predicting anxiety, depression and stress in modern life using machine learning algorithms. Procedia Computer Science, 167, 1258-1267.
15. AlSagri, H. S., & Ykhlef, M. (2020). Machine learning-based approach for depression detection in twitter using content and activity features. IEICE Transactions on Information and Systems, 103(8), 1825-1832.
16. Srimadhur NS, Lalitha S (2020) An End-to-End Model for Detection and Assessment of Depression Levels using Speech. Procedia Computer Science 171:12-21