

## Analysis and Prediction of Mental Health Status using Machine Learning Models

**Dr. Smita Bhanap (Assistant Professor, Fergusson College Pune)<sup>1</sup>**

**Shameek Bhowmick (Student, M.Sc. Data Science, Fergusson College Pune)<sup>2</sup>**

**Sakshi Panhalkar (Student, M.Sc. Data Science, Fergusson College Pune)<sup>3</sup>**

**Abstract:** In this fast-moving world, many people are suffering from psychological health issues like anxiety, depression and stress widely. Numerous factors contribute to mental health issues and affect the overall personality of an individual either positively or negatively. Psychologists and researchers are working on assessment and therapy required for finding the factors affecting mental health of an individual. The need for effective medical health care can be fulfilled by using advanced technologies like machine learning which can provide accurate solutions to mental health problems. This paper presents a study based on mental health status prediction using Supervised Machine Learning classifiers like regression and decision tree. The study involves a survey based on mental health related questionnaire where questions are related to general Psychopathology and other factors. The machine learning models then predict the overall mental health status using relevant factors using the appropriate evaluation metrics.

Keywords: psychological health, supervised machine learning, decision tree, regression.

### Introduction

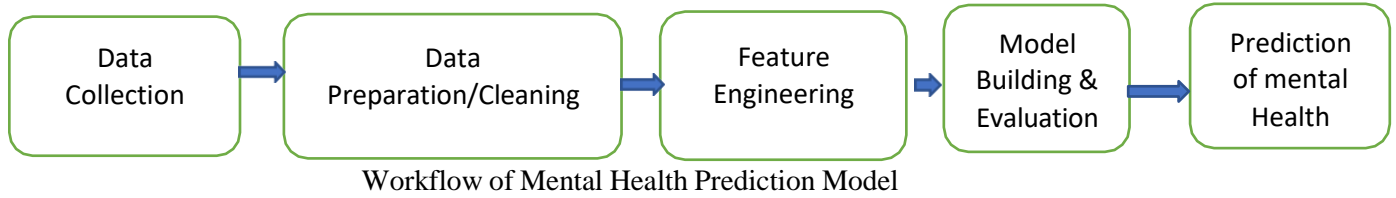
Mental illness is a health problem that undoubtedly impacts emotions, reasoning, and social interaction of a person. These issues have shown that mental illness gives serious consequences across societies and demands new strategies for prevention and intervention [1]. To diagnose a patient's problem, the doctor may ask the patient to fill out a questionnaire. The nature of these questions could be situational and objective which mainly includes depression, anxiety [2]. The World Health Organization (WHO) has observed that depression is the most prevalent mental disorder affecting more than 300 million people worldwide, an appropriate learning algorithm is required for an accurate diagnosis. According to WHO, a healthy person possesses a healthy brain along with physical wellness [3].

### Literature Review

In recent times, many studies and researches contribute to the work where people have been predicting mental health problems like depression and anxiety using the algorithms of machine learning, like decision tree, support vector machine, random forest and convolution neural network [2]. In many studies, python programming has been used for modelling experiments, with the best result among all the classifiers. The authors of [3] shows that the main symptoms of depression from a clinical point of view are loss of memory; lack of concentration; an inability to make decisions which is more dangerous and need immediate and correct diagnosis. Factors like Anxiety and mood disorder were discussed in [3] which can be detected by scanning patient's facial expressions. Authors of [4] discuss various survey methods to predict mental illness which include offline and online surveys using chatbots also [4]. The work in [5] talks about data of individuals working in the industry and predict depression based on multiple work-life and person life-based factors using various machine learning algorithms. The authors of [6] propose ANN based model for mental disorder detection and suggest work life balance also. In [7], authors design and discuss the work related to Mental Health Prediction application, a one stop solution for all the employees where they can check their mental health regularly. WEKA and SPYDER tools can also be used to apply different data mining techniques for the prediction of mental status [8]. The authors of [9] discuss more severe mental diseases like diagnosing bipolar disorder, automatic depression estimation using deep neural networks.

### Methodology

This research study is carried out using the following workflow.



**Data Collection:**

The proposed study has been carried out using a thoroughly prepared survey form which collects human mental state related information based on various parameters. The following parameters and related scale range values play an important role in deciding mental status of a person. Table 1 shows the details of variable parameters and their respective scale values to be considered.

Variables	Scale/ Question	Author	No. of Items/ Questions	Likert Response 1	Likert Response 5	How the final scores are taken
General Psychopathology	SCL-K-9	Klaghofer and Brähler, 2001	9 items	Not at All	Completely	Mean of 9 items
Coping Strategies	Brief COPE	Carver, 1997	14 items ( 2 per strategy)	Never	Very Often	Mean of 2 items per coping strategy
Shyness	Cheek and Buss Shyness Scale	Cheek and Buss, 1981	9 items	Not at all	Completely	Mean of 9 items
Loneliness	NYU Loneliness Scale	Rubenstein and Shaver, 1982	4 items	Strongly Disagree	Strongly Agree	Mean of 4 items
Self-Esteem	Rosenberg Self-Esteem Scale	Rosenberg, 1979	5 items	Strongly Disagree	Strongly Agree	Mean of 5 items
Life-Satisfaction	Satisfaction with Life Scale	Ed Diener et al., 1985	5 items	Strongly Disagree	Strongly Agree	Mean of 5 items

Table 1: Mental Health Status Prediction Chart

The sample set of predicting and target variable are as mentioned below.

PREDICTING VARIABLES	VALUES
GENDER	MALE
AGE	22
PREFERRED COPING STRATEGY	SELF DISTRACTION
SHYNESS	MODERATE
LONLINESS	MODERATE
SELF-ESTEEM	MODERATE
LIFE-SATISFACTION	SATISFIED

TARGET VARIABLE	VALUE
GENERAL PSYCHOPATHOLOGY	HEALTHY

**Data Pre-processing:**

The data collection has been carried out rigorously but still in the excel sheet there are few parameters where data seems to be inadequate and needs cleaning.

Following pre-processing is performed on the excel sheet data generated through the survey form.

- Dropping features which are not required.
- Cleaning erroneous inputs like years in age , gender as male/female only
- Renaming Column as required.
- The data was comprising of responses for ever question of each factor, the mean value for each factor for each person is taken.
- Data dummification was done for sex and location columns.

**Multiple Linear Regression:**

Multiple linear regression is an extension of simple linear regression that involves more than one independent variable. It models the relationship between a dependent variable and two or more independent variables by fitting a linear equation to the observed data.

**Decision Tree Classifier:**

A decision tree classifier is a supervised machine learning algorithm that is used for classification tasks. It is a tree-like structure where each internal node represents a feature or attribute, each branch represents a decision rule based on that feature, and each leaf node represents a class label or a final decision.

**Experimental Results**

After data cleaning part, the data is divided into training(70%) and test (30%) dataset. Then, logistic regression and decision tree classifier are applied to the training data. The experimental results are as shown below.

Algorithm	Train/Test	Accuracy Score
Linear Regression	Train	0.65 (R2 value)
	Test	0.65 (R2 value)
Logistic Regression	Train	0.69
	Test	0.68
Decision Tree Classifier	Train	0.99
	Test	0.95

Table2: Accuracy results of train and test data

As depicted by the accuracy, decision tree is the best model that can be used to predict general psychopathology. The detailed analysis using classification report can be seen in the chart below.

	precision	recall	f1-score	support
Healthy	1.00	0.90	0.95	52
Mild	0.91	1.00	0.95	91
Severe	1.00	0.88	0.94	34
accuracy			0.95	177
macro avg	0.97	0.93	0.95	177
weighted avg	0.95	0.95	0.95	177

Classification Report of Decision Tree classifier

The test dataset is then used for the prediction of mental state of a person which can be seen in the table3 below.

	Age	Sex	Location	Self Distraction	Denial	Venting	Self Blame	Behavioural Disengagement	Acceptance	Active Coping	Shyness	Loneliness	Self-Esteem	Life Satisfaction	Predicted General Psychopathology
Shameek	22	Male	Urban	4	2	3.5	2.5	3	4.5	4.5	2.888889	2	3	3	Mild
Sakshi	22	Female	Urban	4	1	4	2	3	4	3.5	1.888889	1	4.2	3.6	Mild

Table3: Mental Health Status Prediction on test data

**Conclusion & Future Work:**

The survey based on mental health related questionnaire having questions that are related to general Psychopathology and other factors help in predicting mental status. Models using machine learning algorithms like Decision Tree Classifier seemed to be the best out of the lot. In this study only few factors are considered but more other parameters can also be explored and studied in the upcoming enhancements to this work. This study can be further extended to a larger scale data with more attributes and a larger sample size. The mental health prediction models can be useful in following aspects: Research study, workplaces to understand mental status of employees, educational organisations can use the models to help students in analysing their study habits and concentration issues.

**References:**

- 1] Jetli Chung and Jason Teo , “Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges”, Hindawi Applied Computational Intelligence and So Computing Volume 2022,
- 2] Satvik Gurjar, Chetna Patil, Ritesh Suryawanshi, Madhura Adadande, Ashwin Khore, Noshir Tarapore “Mental Health Prediction Using Machine Learning” International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume:09 Issue: 12
- 3] Anu Priyaa, Shruti Garga,\*, Neha Prerna Tiggaa, “Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms” International Conference on Computational Intelligence and Data Science (ICCIDS 2019).
- 4] Vidit Laijawala, Aadesh Achaliya, Hardik Jatta, Vijaya Pinjarkar, “Classification Algorithms based Mental Health Prediction using Data Mining”, International Conference on Communication and Electronics Systems (ICCES 2020)
- 5] Shivangi Yadav, Tanishk Kaim, Shobhit Gupta, Ujjwal Bharti, Prakhar Priyadarshi “Predicting Depression From Routine Survey Data using Machine Learning”, International Conference on Advances in Computing, Communication Control and Networking (ICACCCN) 2022
- 6] Gundeti Sai Manvith, Gunna Shivani Reddy, Bandi Lakshma Reddy “Mental Health Prediction for an individual using Machine Learning” , International Research Journal of Modernization in Engineering Technology and Science Volume:04/Issue:07/July- 2022
- 7] U. Sairam, Santhosh Voruganti, “Mental Health Prediction Using Deep Learning”, IJRASET Volume 10 Issue II Feb 2022
- 8] T. E. Ramya, S.Sindhupriya, “An Effective Approach for Mental Health Prediction Using Machine Learning algorithm” , IJERT, CCICS – 2022
- 9] Ngumimi Karen Iyortsuun , Soo-Hyung Kim ,Min Jhon , Hyung-Jeong Yang Sudarshan Pant , “A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis” , Healthcare 2023, 11, 285. <https://doi.org/10.3390/healthcare11030285>
- 10] Ms. Sumathi M.R., Dr. B. Poorna, “ Prediction of Mental Health Problems Among Children Using Machine Learning Techniques” , International Journal of Advanced Computer Science and Applications, Vol. 7, No. 1, 2016
- 11] Klaghofer, R., and Brähler, E. (2001). Konstruktion und teststatistische Prüfung einer Kurzform der SCL-90-R [Construction and statistical evaluation of a short version of the SCL-90-R]. Z. Klin. Psychol. Psychiatr. Psychother. 49, 115–124.
- 12] Carver, C. S. (1997). You want to measure coping but your protocol’s too long: consider the brief cope. Int. J. Behav. Med. 4, 92–100. doi: 10.1207/s15327558ijbm0401\_6
- 13] Cheek, J. M., & Buss, A. H. (1981). Shyness and Sociability. Journal of Personality and Social Psychology, 41, 330-339. <https://doi.org/10.1037/0022-3514.41.2.330>
- 14] Rubenstein, C., and Shaver, P. R. (1982). “The experience of loneliness,” in Loneliness: A Sourcebook of Current Theory, Research, and Therapy, eds L. A. Peplau and D. Perlman (New York, NY: Wiley-Interscience), 206–223.
- 15] Rosenberg, M. (1979). Conceiving the Self. New York, NY: Basic Books.

- 16] Ed Diener, Robert A. Emmons, Randy J. Larsen & Sharon Griffin (1985) The Satisfaction With Life Scale, *Journal of Personality Assessment*, 49:1, 71-75, DOI: 10.1207/s15327752jpa4901\_13
- 17]