

# Escape The Pain - A Deep Learning Approach for Assisting Depression

<sup>1</sup>Asiya Dalal, <sup>2</sup>Urvita Mhatre, <sup>3</sup>Tejal Thombare, <sup>4</sup>Rachana Dhannawat

<sup>1,2,3</sup>B.Tech Student, <sup>4</sup> Assistant Professor  
Usha Mittal Institute of Technology,  
SNDT Women's University, Mumbai, India

**Abstract :** Depression is a psychological disorder that has influenced various factors, including stress, life-style, physical activity, and physical health. It comes with symptoms such as persistent depression, frustration, and attempts to commit suicide. In health care, it is necessary to accurately predict various life situations. Therefore, the concern for this psychological condition, you should recognize the status of an individual, and guide them. Mental health disorders, as a rule, are accompanied by depression. However, it is not the case that there is a great deal of research, and to predict the situations in order to stop the great depression. Therefore, to find the most accurate model to predict depression, we have been working with a number of models for predicting the risk of major depression. In the field of mental health care and treatment, we use the model to determine the condition of the patient, with the aid of a machine learning algorithm. In this paper, we have analyzed the different machine learning algorithms which can predict depression such as Decision Tree, Extra Trees (Ensemble Technique), Random Forest, Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbour (K-3), Naives Bayes are analyzed to find more accurate model to predict depression. This paper also states that the decision tree has a higher accuracy of 85.75%. The proposed model will take the form of information in order to predict the situations and circumstances that may affect the depression, taking into account the contextual information.

**IndexTerms - Decision Tree Algorithm, Depression risk, Mental health, Machine Learning.**

## I. INTRODUCTION

Depression is a common condition in the world which is affecting many people declared by the World Health Organization (WHO). Depression will challenge our day to day life as it is not similar to short - term emotional reactions such mood swings. The maximum reason of suicide is Depression. As there is a spread of internal and external factors there are people facing mental illness and psychological stress. People might change their moods according to different situations and places also it might change at different times and weather. Most often people who are suffering depression or are at risk of depression, face lack of confidence and enthusiasm. Mostly depression is found in the age of 30 and 40 and it is also found in young adults due to academic stress or interpersonal relations. Therefore, depression is a psychological disorder that is found in a variety of age groups. Since there is negative feedback on the people who suffer such disorders therefore they often hide their mental illness and conditions. And therefore, we find symptoms like lack of enthusiasm, low confidence, etc. According to our study, patients visiting to treat depression are at the age of 50s and 60s, which means that young people are not taking care or any preventive measures in their busy lives. In such situations, there is an increasing demand for rapid treatment and taking measures at the early stage of life. If depression is detected early that it has a high rate to be cured, however, it is likely to come back. In order to eliminate depression and its cause we must take proper care to avoid it and for this it is necessary to prepare in advance and to predict the risk of depression. So we need a model that will be used to gather information on the diet of the people, their diseases, their stress level and other things, to analyse the correlation between depression and other factors on the basis of information collected and analysed. And then we will use this model to predict the depression risk.

## II. LITERATURE SURVEY

Literature survey focuses on studying different methods used for implementing depression detection models and its outcome. The author in this paper[1], proposed a method of time frequency analysis of internet behaviours to predict depression. Here naives bayes, decision tree and logistic regression modelling approaches are used to build models. They used clustering and time-frequency analysis for feature extraction and building classification models for depression detection. They have mentioned that due to time-frequency the performance of the model has improved [1]. Due to the concern of the people suffering from mental health conditions which were leading to major tragedies. [2] The author thought of building a model using the social media data and proposed a deep learning model by studying different approaches to detect depression using social media data. They have studied four models, LSTM, Attention-LSTM, Attention-BiLSTM, X-A BiLSTM. The approach was focused on solving the problem caused by imbalanced data in the real world. Their purpose was to reduce data imbalance, to enhance classification capacity [3]. The author details a variety of methods in order to predict depression[4] that includes the collection of the data sets, by means of personal questions, questions for social media posts, and text messages, which are used in verbal communication, and the expression on his face. In order to identify the mental health situation of the target group and the different machine learning algorithms and classifiers are analyzed, such as Decision Trees, SVM, Naives Bayes', Logistic Regression, and KNN group. In a proposed deep-learning method, you can extract features from data collected from Twitter in order to detect major depression[4]. The author aimed to identify common traits among depressed people[5], with the help of a data set from various social media platforms, where people interact and share their experience, and identify depressed people using various machine learning algorithms. The method used is the detection method by analyzing social media posts, syntactic, and semantic analysis in order to identify a person's emotions and feelings, and to predict the level of depression among the different age groups. The analysis of the algorithm for the detection of emotions and feelings in a tweet, as well as for the detection of suicidal ideation among people in the social network, without checking the validity of the tweets [5]. They have developed a chatbot which takes speech or text as input and provides human-like responses to the text entered by the user[6]. The system determines if the person suffers from depression and also the root problem, or reason behind depression is identified. The responses are tailored to a user and contain article or music suggestions that facilitate research and address the rationale for depression within the user[6]. They presented a new deep multi-tasking recurrent neural

network[7] to deal with depression, how depression classification system is optimized, shared by the two tasks, namely, the only class of metric learning and anomaly ranking. They drew attention to the longitudinal data, (i) temporal changes in the development of the individual and the environment in family that can provide critical signals for mental health disorders, and (ii) that would allow us to predict with certainty what the disorder is, that is, before the disease develops. They have proposed a novel multi-task learning framework and its instantiation MTNet for depression classification. The results show that the MTNet can find in almost every case of depression, when the topics of depression, and the accurate prediction of what to 80% of all cases of depression in the 2 to 4 years prior to occurring depression [4].

### III. DATASET

The dataset is collected by combination of different sources like surveys through google forms. The following are the important features in the dataset which are used for decision making : Age, Gender, Peer Pressure, Depression, Disregard for personal appearance, Self- destructive behavior, Forced career choices, Loss of a loved one, Social media Influence.

### IV. PROPOSED SYSTEM

Machine learning algorithms including Decision Tree, Extra Trees (Ensemble Technique), Random Forest, Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbour (K-3), Naives Bayes are being implemented to design a model for depression prediction.

#### 4.1 Decision Tree

Decision Tree is a Supervised learning technique [1]. It helps in solving Classification and Regression problems, however mainly it is used for fixing Classification problems. It is a tree-based classifier, wherein internal nodes means the functions of a dataset, branches means the decision rules and every leaf node represents the outcome. Decision tree contains leaf nodes and decision nodes. Decision nodes have more than one branch and are used to make any decisions, while Leaf nodes are the output of these selections and that they do now no longer include any similar branches. Decision tree is a graphical illustration for obtaining all of the feasible answers to a problem/choice primarily based totally on given conditions.

#### 4.2 Extra Trees

The Extra trees algorithm generates a large number of unexpected decision trees from the training dataset. In the case of regression, prediction is done by averaging the prediction of decision trees by using majority voting in case of classification. Each of the decision trees of the extra trees forest are put together in the first training sample. Then, for each test at each node in the tree, you are given a random selection of k objects. A series of the functions of each of the decision trees have to pick a unique object to select information that supports a certain mathematical criteria. In order to perform feature selection, with each feature present, the resolution will be adjusted in a descending order according to what is required of them, because each of the features, and the user selects the top k features, according to variety.

#### 4.3 Random Forest

Random forest is a supervised learning technique used for classification problems in machine learning [5]. It supports the concept of ensemble learning. Ensemble learning is the process of mixing multiple classifiers. Instead of taking predictions of a decision tree, random forest takes the predictions from every tree, and supports one of the largest votes of predictions, as well as predicts the final outcome. A large number of trees and forests, leading to greater accuracy and to prevent overfitting.

#### 4.4 Logistic Regression

One of the algorithms that are used for the classification of supervised learning is logistic regression [5]. For predicting the probability of a target variable, logistic regression is used. The nature of target or dependent variable could only have two possible classes. It can also be used to classify the observations, with the help of different types of data. It is easy to determine the most efficient and effective solutions for the variables used for the classification. These are the easiest of the ML algorithms that could be used in a variety of classification tasks, such as spam detection, the prediction of diabetes, cancer detection and so on.

#### 4.5 Support Vector Machine

Support Vector Machine (SVM) is a machine learning algorithm [2]. It is used mostly in classification rather than regression. The objective of this algorithm is to plot each data item in n-dimensional space where each feature assesses a particular coordinate. Then the classification is done by finding the hyperplane that separates the two groups. Support vectors are the data points which are closest to the hyperplane. These data points help to define separating lines. To find the maximum marginal hyperplane (MMH), SVM separates the datasets into classes. SVM creates hyperplanes iteratively that separates the classes in the best way and selects the hyperplane correctly.

#### 4.6 K-Nearest Neighbour

The K-nearest neighbour algorithm is a supervised Learning technique. It implies the similarity between a new data, and available data, and adds a new set of events/data in the category, which is the most similar to the existing categories. It will store all of the available data and use it to classify new data points. It is mostly used for classification problems. This is a nonparametric algorithm. In this algorithm, the information is collected during the training phase and the dataset is stored and when it gets new data, then the data fall into the category that is much alike to the new data.

#### 4.7 Naives Bayes

This classification method is based on Bayes' theorem, with the predictor hypothesis. In simple terms, a Naives Bayes [2] group assumes the existence of some of the features of a category, unrelated to other features that will be present. The Naive Bayes model is easy to set up and is particularly useful for very large data sets. Along with simplicity, Naive Bayes is considered superior to even the very complex methods.

## V. EXPERIMENTAL PROCEDURES

We had obtained the consolidated data of the people from different files for analysing. Then we searched(treated) the missing values, outliers, duplicates if any. Used data visualization to get the insights of the data. The year-wise suicide count data will take us to a broad overview about the different features of the persons such as height, weight, age, professional profile, income, etc. affecting the suicidal count in India.

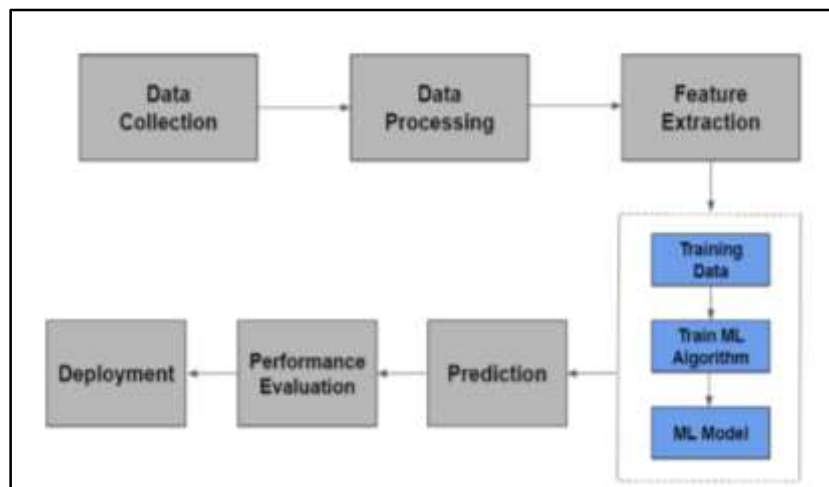


Fig. 1. Methodology

### 5.1 Data Collection

In this process, we had collected data from different aspects of a person's life and used that data as a dataset. The collected data was in different forms (such as .txt, .xlsx, .csv), so all the data was converted into standard format that is in .csv format. In various files and folders, the data was scattered, so a single csv file was generated by merging the data. Then that data was used to train and test the system.

### 5.2 Data Cleaning

There might be incomplete or unknown values in the collected data hence we have to remove those. In this process, we had removed missing fields(values) and considered only the common attributes. To remove the missing values, `df.isna().sum()` function of pandas in python was used which returns the number of missing fields in each column and after that the missing fields were dropped using `dropna()` function of pandas in python.

### 5.3 Training

Now we will use the collected data sets to train the machine using a machine learning algorithm. After data collection and data cleaning, the pre-processed data was used to train the model using different machine learning algorithms.

### 5.4 Testing

After training the model we had tested the model using the same data set.

## VI. DATA ANALYSIS

The majority of suicides are associated with depression, and among young people the leading cause of death is suicide. Even though most of the people who have depression do not die by suicide, having depression increases the risk for suicide compared to those without depression. The risk of death from suicide, which may be partly related to the severity of the depression. So keeping this in mind, the Hospital Anxiety and Depression Scale (HADS) score can be considered directly proportional to the possibility of suicide. Various factors affect the rate of depression as we have mentioned above in the features of the data set. Using exploratory data analysis skills we can find inferences of various factors affecting depression and suicides in turn.

## VII. RESULTS AND DISCUSSIONS

Different machine learning algorithms such as Decision Tree, Random Forest, Logistic Regression, Extra Trees (Ensemble Technique), Naives Bayes, Support Vector Machine (SVM), K-Nearest Neighbour (K-3) performances were compared. In which decision tree gives us the highest accuracy.

Table 1 Model Accuracy

Model Name	Accuracy Score(%)
Decision Tree	85.75
Extra Trees (Ensemble Technique)	85.73
Random Forest	85.74
Logistic Regression	73.94

Support Vector Machine (SVM)	84.68
K-Nearest Neighbour (K-3)	82.35
Naives Bayes	66.19

After implementing these machine learning algorithms and comparing their accuracy after training and testing the data, Decision Tree was the best fit for depression prediction.

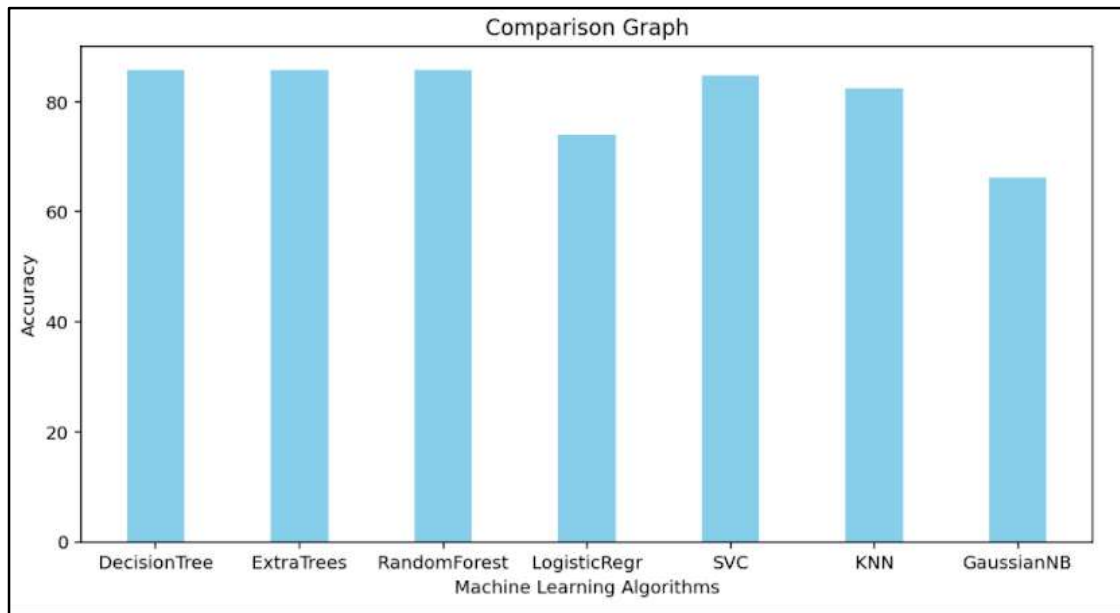


Fig. 2. Comparison of Accuracy of different Algorithms

We have created a flask web application which we have linked with a simple html form and according to our trained model whenever someone clicks on the submit button of html form it will predict the depression risk of a person.

Fig. 3. Data collection from Patient Case 1 : Low Risk

Fig. 4. Predicted Output Case 1



Escape the Pain

Age: 22

BMI: 30

Gender: Female

Sleeping Difficulty: Several Days

Suicidal Thoughts: No

Smoking/Drinking: No

Do you Exercise Regularly: Sometimes

Feeling Sad/Empty: No

Lost Of Interest: No

Family Health Issues: No

Doctor Consultation: Several Times

Submit

Fig. 5. Data collection from Patient Case 2 : High Risk

Escape The Pain

Your risk of Depression has been predicted below!

Danger! You are at high risk of Depression.

Fig. 6. Predicted Output Case 2

## VIII. CONCLUSION

This model is aimed toward predicting depression at an early stage, with a high degree of accuracy. It has been trying to predict depression disorders an individual has. Therefore, other diseases that are rooted in the depression of the early stages, can also be prevented from the initial stage itself. Here the stable machine learning algorithms are used and are compared to test and find the classifier with high accuracy to predict in which five algorithms have accuracy more than 82% and decision tree gives us the highest accuracy of 85.75%. Therefore, Decision Tree was the best fit for depression prediction.

## REFERENCES

- [1] C. Zhu, B. Li, A. Li and T. Zhu, "Predicting Depression from Internet Behaviors by Time-Frequency Features", 2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI), pp. 383-390, (2016)
- [2] P. V. Narayanrao and P. Lalitha Surya Kumari, "Analysis of Machine Learning Algorithms for Predicting Depression", 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), pp. 1-4, (2020)
- [3] Q. Cong, Z. Feng, F. Li, Y. Xiang, G. Rao and C. Tao, "X-A-BiLSTM: a Deep Learning Approach for Depression Detection in Imbalanced Data", 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1624-1627, (2018)
- [4] Pang, Guansong & Pham, Ngoc & Baker, Emma & Bentley, Rebecca & Hengel, Anton, "Deep Multi-task Learning for Depression Detection and Prediction in Longitudinal Data", CoRR, Vol.abs/2012.02950 (2020)
- [5] D. Ramalingan, V. Sharma and P. Zar, "Study of Depression Analysis using Machine Learning Techniques", International Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol. 8, (2019)
- [6] Shweta Oak, "Depression Detection and Analysis", AAI, pp. 710-713, (2017)