# Depression Prediction using Emotion Detection and Text Mining to prevent Suicide using Machine Learning

**Guided By- Prof. N. R. Zinzurke**
Kavita S. Munji
Student of Computer Engineering
Savitribai Phule Pune University
Pune, India

**Abhishek B. Nimbalkar**
Student of Computer Engineering
Savitribai Phule Pune University
Pune, India

**Nisha B. Pawar**
Student of Computer Engineering
Savitribai Phule Pune University
Pune, India

**Supriya Sanjay Yenare**
Student of Computer Engineering
Savitribai Phule Pune University
Pune, India

***Abstract*—**

**Suicide is one of the most serious social health issues that exists in today's culture. Suicidal ideation, also known as suicidal thoughts, refers to people's plans to commit suicide. It can be used as a suicide risk measure. India is among the top countries among in the world to have annual suicide rate. Social networks have been developed as a first rate factor for its users to communicate with their interested buddies and proportion their captions, photos, and videos reflecting their moods, emotions and sentiments. To increase and put in force a version which takes a facial expression images as an enter and symptoms. On the basis of that it predicts the repute of that patient whether or not he/she has been detected or now not detected for depressed. We can train version using photographs & will use it for prediction. Image captioning can be accomplished after prediction for higher visualization of report. We will also use text mining (NLP) technique to predict melancholy the usage of signs furnished with the aid of person.**

**At final we are able to make final choice primarily based on above two techniques. To generate detailed dashboard of user disease status and to design webapp for above system. We will use CNN algorithm for speed up detection of depressed character instances and approach to become aware of high quality answers of mental health troubles. We suggest system learning method as an efficient and scalable technique. We document an implementation of the proposed method. We've evaluated the efficiency of our proposed technique the usage of a set of various psycholinguistic features. We show that our proposed method can extensively improve the accuracy and category blunders price. Key Words: Emotion Recognition, Depression, Convolutional Neural Networks, Text processing, Image processing, Sentiment analysis**

*Keywords: Security, Reliability, Data Integrity, Block chain, health care, brain tumor.*

## I. Introduction

In the Indian way, suicide is a big problem. Suicide kills over lakh people (100,000) in our country each year. The suicide rate has risen from 7.9 to 10.3 per 100,000 over the past two decades. There are a variety of suicide rates in the world. Kerala, Karnataka, Andhra Pradesh, and Tamil Nadu are among the southern regions with the highest suicide rates. For the past two decades, this trend has continued. High suicide rates in the southern states may be explained by higher education, a stronger reporting system, lower external violence, a higher socioeconomic status, and higher ambitions. The number of suicides in India has risen to 230,314 in 2016. Suicide was the leading cause of death in both 15-29 and 15-39 age groups. Every year, about 800,000 people commit suicide worldwide, of which 135,000 (17%) are Indian citizens, accounting for 17.5% of the world's population. "Suspended" suicide (53.6%), "poisoning" (25.8%), "drowning" (5.2%), and "immersion" (3.8%) were the most common forms of suicide throughout the year, according to the report. According to a new study by the World Health Organization (WHO), India had the highest suicide rate in the Southeast Asian region in 2016. For the past three years, India's official statistics, showing the number and causes of suicide in the country. , have not yet been identified, disrupting suicide prevention strategies and efforts to enforce WHO recommendations in the region. The study used data from the 2016 WHO Global Health Estimates to present suicide rates nationally and regionally. India is a region of Southeast Asia and a region of the Lower Middle-Income countries in terms of region and revenue. The suicide rate in India (16.5) was higher than in the surrounding regions (13.4) and in the group of immigrants (16.5) (11.4).

## II. Literature Survey

There is a growing body of research on stress factors [9 - 12]. Choudhury et al. [13] suggest that depression is a true measure of personal and social well-being. A large number of people suffer the negative effects of depression, but only a small percentage receive appropriate care each year. They also looked at the possibility of using social media to identify and evaluate any signs of severe depression in humans. They rated ethical credits associated with social interactions, emotions, dialect and semantic forms, self-explanatory system descriptions, and stress-relieving drug notes in their webbased writing. Choudhury et al. [14] saw online communication as a promising public health tool, focusing on the use of Twitter to build predictable models on the effect of childbirth on the behavior and behavior of young mothers. They used a Twitter post to track 376 maternal changes in relation to communication, emotions, and information. [15] It has been found that Twitter is increasingly being investigated as a tool for diagnosing psychological problems. Depression and social ills are examples of poor mental health for many people It was discovered during their research that it is possible to determine the level of anxiety among people who want to commit suicide. Using both human codes and machine learning algorithm, we were able to find similar tweets. Organized computer class Several studies have shown that the optimal use of user-generated content (UGC) will help determine people's psychological well-being. Aldarwish and Ahmad [17], for example, have found that the use of Social Network Sites (SNS) is increasing these days, especially among younger generations. Clients may share their wishes and feelings on social media because they are available. Using emotions, psycholinguistic processes, and drug titles extracted from posts produced by people from these groups, Nguyen et al. [20] used machine learning and mathematical methods to distinguish online messages between stress and control groups. Park et al. [21] look at people's attitudes and behaviors about web-based social networking sites to see if they are depressed. They organized slightly built one-person meetings with 14 active Twitter users, half of whom were depressed and the other was absent. Alternatively, they look at a number of ways to address potential communication strategies that may better suit depressed users and include information to help depressed users deal with their problems through web-based social networking sites [22]. Holleran [9] found the first evidence that depression has a significant impact on the global epidemic of Facebook. Wang et al. [19] and Shen et al. [25] looked at various stress-related factors and developed a multimodal stress model to identify depressed users. While some of the research mentioned above looked at emotional processes, transient processes, and language style to diagnose depression, current literature has the following errors: SVM, KNN, Decision Tree, and Ensemble have all been used independently in a few studies. There are no known studies that have investigated differences in strategic outcomes using both methods in the same database. Not many studies have used the above machine learning strategies to detect depression using Facebook data. To correct the above errors, we are trying to diagnose depression from Facebook comments on this page; we also increase access to stress measures based on social media by explaining the various aspects of Facebook user comments. We used machine learning methods to find depressed people using certain testsof this fell CNN process was likewise planned.

## III. A.Problem Definition:

Designing a system that involves the removal of facial features, as well as the detection of pressure based on facial expressions using the Convolutional Neural Network (CNN) algorithm and separating positive and negative emotions and receiving pressure based on the normal limit value.

## IV. Proposed Methodology:

Face of the subject is captured using the camera module. This detected face is processed and the emotions are classified as either positive or negative emotions. The detected image is processed to identify the face of the subject using Convolutional Neural Network (CNN) algorithm.

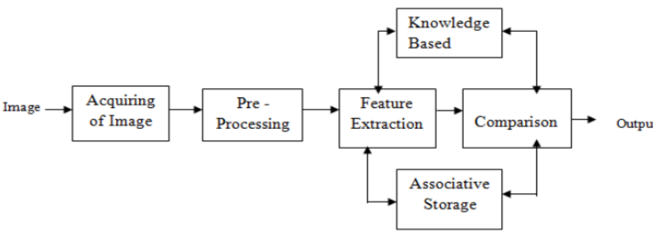



Fig.1 Methodology Of the system

This is plotted and an increase in the negative emotion can be inferred as increase in stress.

- Face Detection

Face Detection is the first and essential step for processing, and it is used to detect faces in the images. A facial

detection system uses biometrics to map facial features from a photograph or video. It compares the information with a

database of known faces to find a match. Face detection systems use computer algorithms to pick out specific,

distinctive details about a person's face.

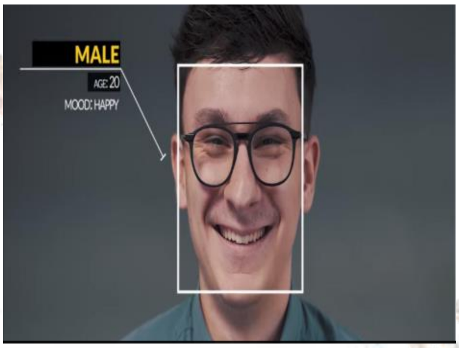



Fig. 2 face detection

These details, such as distance between the eyes or shape of the chin, are then converted into a mathematical epresentation and compared to data on other faces collected in a face database.

- Emotion Detection

Emotion detection is used to analyze basic facial expression of human. Emotion recognition system is constructed, including face detection, feature extraction and facial expression classification. The process of splitting an image into multiple parts is known as segmentation. It creates various sets of pixels within the same image. Segmenting an image makes it easier for us to further analyze and extract meaningful information from it. It is also described as “The process of labeling each pixel in an image such that they share the same characteristics”. The process results in pixels sharing a common property.

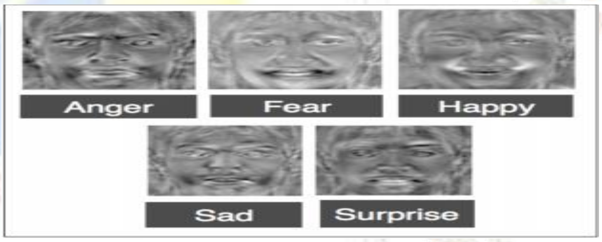



Fig 3 Emotion Detection

- Feature Extraction

Facial feature extraction is the process of extracting face component features like eyes, nose, mouth, etc. from human face image.

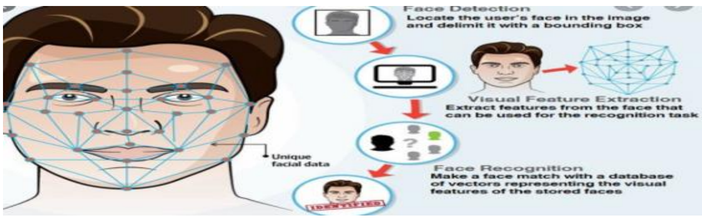



Fig. 4. Feature Extraction

Facial feature extraction is very much important for the initialization of processing techniques like face tracking, facial expression recognition or face recognition.

- Emotion Recognition

The emotions are to be extracted from the detected face. The image that is captured from the camera module, contains the facial features. The detected face is pre-processed (i.e.) cropped and resized. The detectors defined prior can be utilized to identify the emotion and sort them. It must be noted that viola-jones algorithm uses adaboost algorithm with cascading classifier, wherein a series of weak classifier's classification with a satisfactory threshold is combined to give an acceptable outcome.

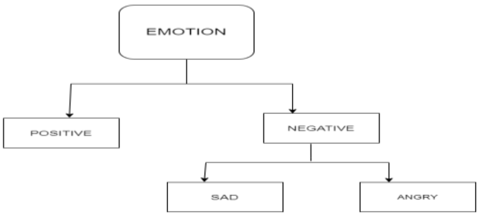



Fig.5 Emotion Recognition

- **Mathematical Model**

Receive input data, process the information, and generate output

Step 1: Load the input images in a variable (say X)

Step 2: Define (randomly initialize) a filter matrix. Images are convolved with the filter Z1 = X * f

Step 3: Apply the Relu activation function on the result

A = Relu(Z1)nf

Step 4: Define (randomly initialize) weight and bias matrix. Apply linear transformation on the values

Z2 = WT.A + b

Step 5: Apply the Relu function on the data. This will be the final output

O = Relu(Z2)

- Algorithm Details

1) Algorithm 1/Pseudo Code

- Image Processing:

In computer science, image processing is the use of computer algorithms to perform image processing on digital images. We used image processing for detecting the faces from camera and to capture emotions on the detected images.

Steps for Image Detection :

Step 1:

Confirm the upper limit of the number of faces to be detected.

Step 2:

Adjust the scaling of the images according to the Device's Camera.

Step3:

Give access of the device's camera (to on and off) and pass the camera port as input to OpenCV library's VideoCapture method.

Step4 : Confirm the frequency of frames needed from the video and capture them within adjusted intervals.

2) Algorithm 2/Pseudo Code

Deep Convolutional Neural Network (DCNN):

Input: Test Dataset which contains various test instances TestDBLits [], Train dataset which is build by training phase TrainDBLits[] , Threshold Th.

Output: HashMap ≤class label, SimilarityWeight ≥all instances which weight violates the threshold score.

Step 1: For each read each test instances using below equation

$$testFeature(m) = \sum_{m=1}^{n} (.\ featureSet[A[i] \ldots\ldots A[n] \leftarrow \text{TestDBLits})$$

Step 2 : extract each feature as a hot vector or input neuron from testFeature(m) using below equation.

$$\text{Extracted\_FeatureSetx}[t\ldots\ldots n] = \sum_{x=1}^{n}(t) \leftarrow testFeature\,(m)$$

Extracted FeatureSetx[t] contains the feature vector of respective domain.

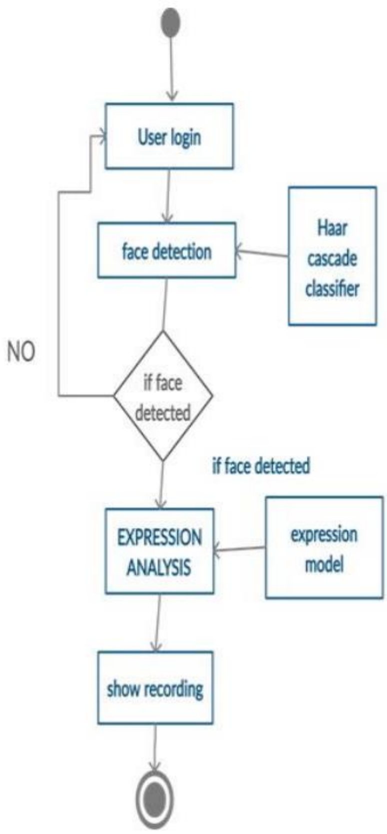
Step 3: create the number of Convolutional

For each read each train instances using below equation.

**Algorithm Used:**

1. CNN(Image Processing & DL)
2. NLP(Text mining)

## V. Flow Chart of the System




## VI. Advantages

The doctor identify the disease earlier and improve patient outcomes drastically. Today, advanced Medical Imaging offers numerous benefits to both the healthcare providers and the patients. CNN is the best approach for medical image processing to find accurate and quick result. Following some advantages of our system is helpful for:

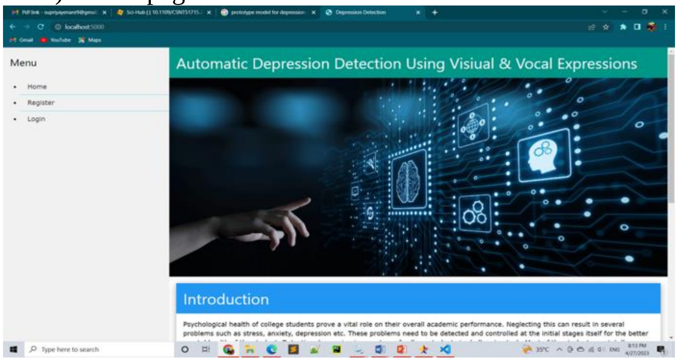
1. Better Diagnosis
2. Complicated Surgeries
3. Affordable Health Care Costs
4. Safe & effective
5. File-sharing Ecosystem & Data Privacy
    - High Accuracy.
    - Less efficient.

## VII. Applications

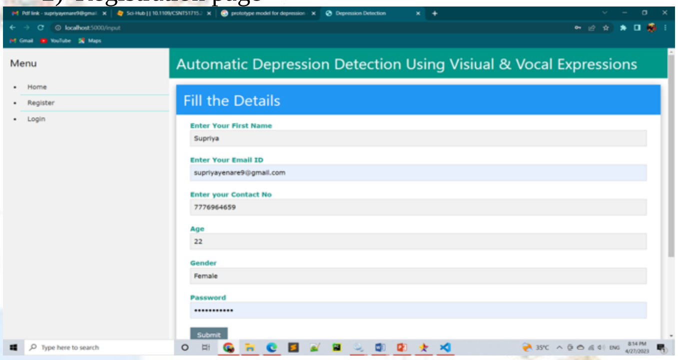
- Leaf Disease Detection.
- Medical image processing

## VIII. Implementation

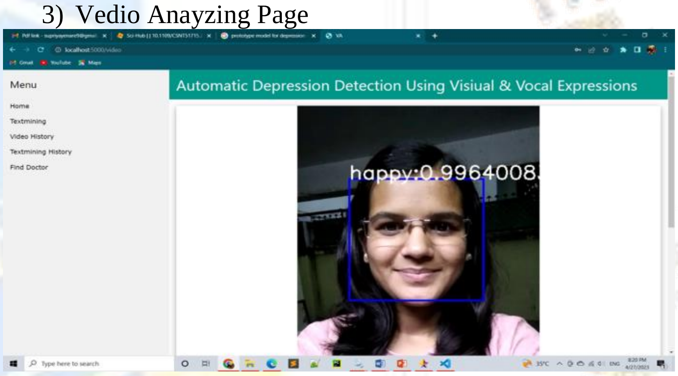
1) Home page



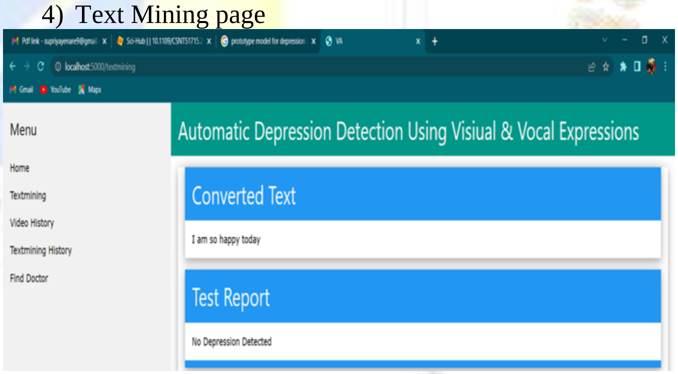

2) Registration page




3) Vedio Anayzing Page




4) Text Mining page




## IX. Conclusion

The prediction was successful compared to predicting test data from the same database used to train variants. However, the predictor remains poor in finding a statement associated with contempt. This may be due to a combination of lack of training and test images that clearly show contempt, poor labeling of previous data training, and internal difficulties in identifying contempt. The class divider also fails to predict the sensitivity of the test data to not only one of the seven key expressions, as they are not trained in other expressions. Future work should include improving the strength of class dividers by adding more training images from different data sets, investigating more accurate detection methods that still maintain mathematical performance, and considering classification of friendly and complex expressions.

## IX.References

[1]. Scott J. Social network analysis. Thousand Oaks: Sage; 2017.

[2]. Serrat O. Social network analysis. In: Knowledge solutions. Singapore: Springer; 2017. p. 39–43

[3]. Mikal J, Hurst S, Conway M. Investigating patient attitudes towards the use of social media data to augment depression diagnosis and treatment: a qualitative study. In: Proceedings of the fourth workshop on computational linguistics and clinical psychology—from linguistic signal to clinical reality. 2017.

[4]. Conway M, O'Connor D. Social media, big data, and mental health: current advances and ethical implications. Curr Opin Psychol. 2016;9:77–82.

[5]. Ofek N, et al. Sentiment analysis in transcribed utterances. In: Pacific-Asia conference on knowledge discovery and data mining. 2015. Cham: Springer.

[6]. Yang Y, et al. User interest and social influence based emotion prediction for individuals. In: Proceedings of the 21st ACM international conference on Multimedia. 2013. New York: ACM.

[7]. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol. 2010;29(1):24–54.

[8]. Pennebaker JW, Francis ME, Booth RJ. Linguistic inquiry and word count: LIWC 2001, vol. 71. Mahway: Lawrence Erlbaum Associates; 2001. p. 2001.

[9]. Holleran SE. The early detection of depression from social networking sites. Tucson: The University of Arizona; 2010.

[10]. Greenberg LS. Emotion-focused therapy of depression. Per Centered Exp Psychother. 2017;16(1):106–17.

[11]. Haberler G. Prosperity and depression: a theoretical analysis of cyclical movements. London: Routledge; 2017.

[12]. Guntuku SC, et al. Detecting depression and mental illness on social media: an integrative review. Curr Opin Behav Sci. 2017;18:43–9.

[13]. De Choudhury M, et al. Predicting depression via social Media. In: ICWSM, vol. 13. 2013. p. 1–10.

[14]. De Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York: ACM; 2013.

[15]. O'Dea B, et al. Detecting suicidality on Twitter. Internet Interv. 2015;2(2):183–8. Signs Symptoms Behaviour Not going out any longer Not completing things at work Not doing regular charming exercises Unfit to focus Feelings Overwhelmed Blameworthy Irritate Disappointed Unlucky Worried Thoughts He is winner It's my pleasure Nothing good ever happens to me He was unlucky Life is not the bed of roses He would not be able to work without me Physical Tired Illness Headaches Depression problem Misfortune Islam et al. Health Inf Sci Syst (2018).

[16]. Zhang L, et al. Using linguistic features to estimate suicide probability of Chinese microblog users. In: International conference on human centered computing. Berlin: Springer; 2014.

[17]. Aldarwish MM, Ahmad HF. Predicting depression levels using social media posts. In: 2017 IEEE 13th international Symposium on Autonomous decentralized system (ISADS). 2017.

[18]. Zhou J, et al. Measuring emotion bifurcation points for individuals in social media. In: 2016 49th Hawaii international conference on system sciences (HICSS). 2016. Koloa: IEEE.

[19]. Wang X, et al. A depression detection model based on sentiment analysis in micro-blog social network. In: Trends and applications in knowledge discovery and data mining (PAKDD). 2013.

[20]. Nguyen T, et al. Affective and content analysis of online depression communities. IEEE Trans Affect Comput. 2014;5(3):217–26. [21]. Park M, McDonald DW, Cha M. Perception differences between the depressed and non-depressed users in Twitter. In: ICWSM, vol. 9. 2013. p. 217–226.

[22]. Wee J, et al. The influence of depression and personality on social networking. Comput Hum Behav. 2017;74:45– 52.

[23]. Bachrach Y, et al. Personality and patterns of Facebook usage. In: Proceedings of the 4th annual ACM web science conference. 2012. New York: ACM.

[24]. Ortigosa A, Martín JM, Carro RM. Sentiment analysis in Facebook and its application to e-learning. Comput Hum Behav. 2014;31:527–4