

Depression Detection using Machine Learning

Vedant Bhamre¹, Prof.R.T.Umbare², Danish Tamboli³, Trushant Jadhav⁴ and Pranav Kurle⁵

¹⁻⁵JSPM's Rajarshi Shahu College of Engineering, Pune, India

Email: vedantbhamre17@gmail.com, rtumbare_it@jspmrscoe.edu.in

dainshtamboli42@gmail.com, jadhavtrushant50@gmail.com, pranavkurle07@gmail.com

Abstract—Suicide is the most significant social health issues our society is now dealing with. Suicidal ideation, also referred to as suicidal thoughts, describes people's intentions to end one's life. It can be applied to determine the suicide risk. India has one of the highest annual suicide rates in the world. Social networks have been created as a fantastic communication tool for users with people who share their interests and post comments, images, and videos that convey their emotions. to create and put into practice a model that incorporates facial expressions as symptoms and indications. That information can be utilized to predict a patient's reputation and establish whether or not they have been given a depression diagnosis. With the use of images, we can train a model that we'll use to provide predictions. For improved report visualization, image captioning can be finished after prediction. Additionally, we'll use text mining (NLP) technologies to predict melancholy based on clues given by people.

Index Terms— Convolutional Neural Network, Emotions, and Suicide Rate.

I. INTRODUCTION

Suicide is a big issue in Indian culture. In our nation, suicide claimed the lives of around one lakh (100,000) individuals annually. In the last 20 years, the suicide rate has increased from 7.9 to 10.3 per 100,000 people. Suicide rates vary widely throughout the world. Tamil Nadu, Kerala, Karnataka, and Andhra Pradesh are some of the southern states with the highest suicide rates. Over the past 20 years, this inconsistent trend has continued. The greater suicide rates in the southern states may be attributed to factors such as better literacy rates, a more efficient reporting system, a lack of external hostility, higher social status, and higher aspirations. In India, 230,314 people committed suicide in 2016. Suicide was the major cause of death in the age group of 15-29 and 15-39. Each year, more than 800,000 people worldwide end life, with 135,000 (17%) of them being Indian nationals, who account for 17.5% of the global population. According to the report, the most common methods of suicide during "Hanging" (53.6%), "eating poison" (25.8%), "drowning" (5.2%), and "self-immolation" (3.8%) were the most common crimes throughout the year.

In South-East Asia in 2016, India has the highest suicide rate, latest World Health Organization (WHO) report claims. It has been difficult to implement WHO recommendations in this area and difficult for suicide prevention programmers since India's own official statistics, which display the number and reasons of suicides in the nation, have been unavailable for the past three years. The study presented suicide rates for different nations and geographical regions using data from the 2016 WHO Global Health Estimates. When nations are categorized according to their geography and income, India is a member of the Lower Middle-Income Countries category and the South-East Asia area. India has a higher suicide rate (13.4) than its region and socioeconomic level. (16.5). (16.5). (11.4).

II. LITERATURE REVIEW

The body of research on the symptoms of depression is expanding [9–12]. Depression is a real condition, claim Choudhury et al. [13] indicator of both an individual's and society's level of wellbeing. Even though many people experience the harmful impacts of depression, only a small proportion of them obtain adequate care each year. They also investigated the potential for identifying and evaluating any indications of severe depression in people utilizing social media. They examined behavioural credits connected to social contact, emotion, Through their online social networking posts, people may learn about dialect and semantic types, the meaning of the self-system, and information on depressive drugs.

With regard to the effects of delivery on new moms' behaviour and disposition, Choudhury et al.'s [14] research concentrated on the usage of Twitter to develop prediction models. They saw internet networking as an effective instrument for promoting public health. They monitored 376 moms' postpartum changes in terms of social interaction, moods, and information using tweets from Twitter. [15] I discovered that Twitter is getting more attention as a method for spotting psychiatric problems. Poor mental health in the general population includes, for instance, depression and social dysfunction. They observed during their analysis that it is possible to gauge how anxious suicidal individuals are. We discovered related tweets by using a machine learning algorithm as well as human coders. an algorithmic computer classifier According to numerous studies, appropriately utilizing UGC can help determine a person's psychological health. AL For instance, Darwish and Ahmad [17] discovered that younger generations are increasingly using social networking sites (SNS).

Due to the accessibility of social media platforms, users may express their wants and sentiments there. Nguyen et al. [20] used mood, psycholinguistic processes, and drug subjects excluded from the postings made by members of these groups to distinguish between online communications from depression and control groups. Park et al. [21] studied people's attitudes and actions towards online web-based social networking to ascertain whether or not individuals were discouraged. They organized semi-structured one-on-one meetings with the 14 active Twitter users, half of whom were discouraged and the other half weren't. They also looked at a few strategic implications for future social networks that would better serve users who are sad and contain information to help discouraged people resolve their problems via online web-based social networking [22]. By searching Facebook, Holleran [9] found early evidence that depression significantly contributes to the entire global burden of illnesses. In order to detect depressed users, A multimodal depressed model was developed by Wang et al. [19] and Shen et al. [25] after they looked at a variety of depression-related features.

The existing body of literature has the following shortcomings despite some of the research listed above looking at emotional processes, temporal processes, and linguistic style to diagnose depression:

Many research have separately employed SVM, KNN, Decision Trees, and Ensemble. No well-known research that used both methodologies on the same dataset to compare the variations in technique-based findings exist. The application of the aforementioned machine learning methods to identify depression using Facebook data has not been the subject of any significant research. In this study, we aim to diagnose depression using Facebook comments and, by elucidating the many traits of Facebook user comments, we extend the applicability of social media-based depression measures. For the purpose of identifying those who are depressed.

III. MOTIVATION

- Although many people experience the harmful effects of depression, only a small percentage of them obtain adequate therapy each year.
- Investigated the potential for using online networking to identify and assess any signs that a person might be experiencing a significant depressive disorder.
- They assessed behavioural factors linked with social engagement, emotion, linguistic and semantic patterns, sense of the self-system, and allusions to antidepressant drugs through their online social networking postings.

For this Machine Learning (ML), data-driven algorithms can be utilized to validate existing methods and help make potential new decisions.

IV. PROBLEM DEFINITION

The project's objective is to address the problem of early depression detection in humans using aural and visual inputs. Millions of individuals throughout the world suffer from depression, a prevalent mental condition. But because mental health issues are stigmatized or poorly understood, they frequently go unnoticed and untreated. Early detection and treatment have a substantial impact on both the chances of recovery and the severity of

symptoms. Convolutional Neural Network (CNN) technology is used to extract facial features, detect stress by classifying positive and negative emotions expressed on the face, and detect stress based on a predetermined threshold value.

V. PROPOSED METHODOLOGY

In this project, the camera is used to collect faces. The emotions on this detected face are analyzed and categorized as either positive or negative. The Convolutional Neural Network (CNN) technique is used to process the detected image and identify the subject's face.

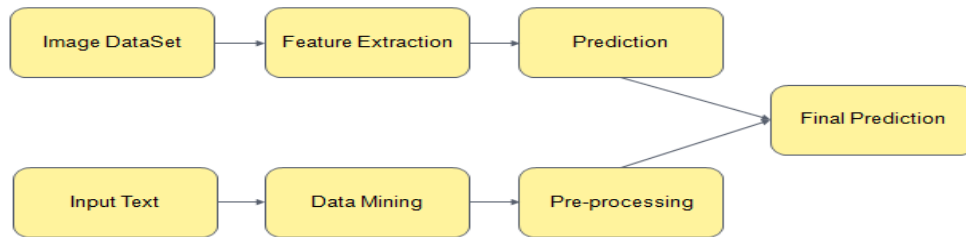


Figure 1: Methodology

The camera module is used to capture the subject's face. The emotions on this detected face are analyzed and categorized as either positive or negative. The Convolutional Neural Network (CNN) technique is used to process the detected image and identify the subject's face.

Algorithms:

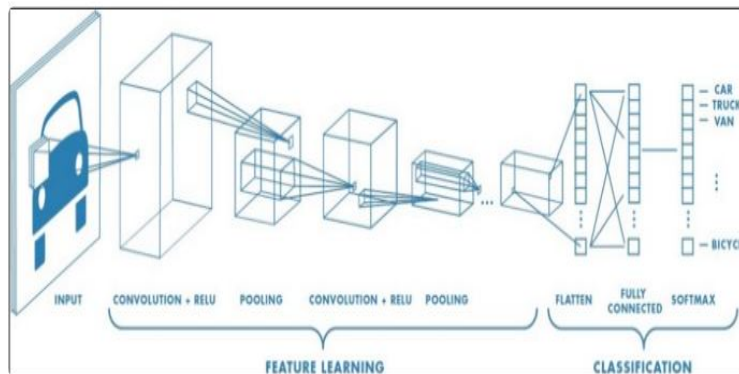
- A. CNN (Image Processing, Convolutional Neural Network)
- B. NLP

Features:

- 1. Depression Detection using Facial Expression (Emotions detection).
- 2. More accurate result produce.

A. Convolutional Neural Network (CNN) algorithm-

One of the primary neural network architectures used to distinguish and categorise images is convolutional neural networks (ConvNets or CNNs). Face recognition, item identification, and other uses are a few of the applications for CNNs. After processing an input image, CNN's image classifiers categorise it into one of several groups. (Eg., Dog, Cat, Tiger, Lion). Computers interpret an input image as a collection of pixels depending on the image resolution. Depending on the image resolution, it will see $h \times w \times d$ (h = Height, w = Width, d = Dimension). For instance, a picture with a $6 \times 6 \times 3$ array of RGB and a $4 \times 4 \times 1$ array of grayscale matrix. Each input picture will formally through a series of convolution layers with filters (Kernels), Pooling, fully connected layers (FC), and the Softmax function with probabilistic values between 0 and 1 in order to categorise an object. To train and test CNN deep learning models, this is done. The flowchart below illustrates how CNN analyses an input image and classifies the items according to values.



B. Natural language Processing (NLP)

The goal of artificial intelligence has always been to create machines that can simulate the capabilities and operations of the human mind. The development of language is regarded as one of humanity's greatest accomplishments that has sped up the evolution of the species. It is not unexpected that Natural Language Processing (NLP), a branch of artificial intelligence, is putting a lot of work into incorporating language.

Why NLP –

Over 50 corpora and lexical resources can be accessed through NLTK's user-friendly interfaces. The tool has the fundamental features needed for nearly all Python operations involving natural language processing. NLTK uses 'Naive Bayes classification algorithms. Typical methods for obtaining data from texts include:

- Named Entity Recognition
- Sentiment Analysis
- Text Briefing
- Aspect Mining
- Subject Modeling

1. Named Entity Recognition

identifies things in the text such persons, places, organisations, dates, etc. NER often uses supervised models and grammatical rules as its foundation. However, there are NER platforms with pre-trained and integrated NER models, like open NLP.

2. Sentiment Analysis

- Especially beneficial in circumstances where individuals may voice their opinions and critiques, as through consumer surveys, reviews, and social media comments
- Both supervised and unsupervised methods can be used for sentiment analysis.
- Naive Bayes is the most often used supervised model in sentiment analysis.
- A model must be built using a training corpus containing sentiment labels. is then employed to determine the sentiment after training.

1. Text Briefing - Text summarizing is frequently employed in situations involving news items and academic research publications. Two general methods for summarizing a text are extraction and abstraction. Extraction techniques result in a summary by excluding certain portions of the text. By creating new language that conveys the key concepts of the original content, abstraction techniques create summaries.

2. Aspect mining identifies the many elements of the text, and when combined with sentiment analysis, it pulls all the information from the text.

VI. RESULT AND DISCUSSION

One of the most important social health issues is suicide. issues our society is now dealing with Suicidal ideation, commonly referred to as suicidal thoughts, describes people's intentions to end their lives. India has one of the highest yearly suicide rates in the world. The creation of social networks has made it possible for individuals to engage with people they are interested in friends and share remarks, photographs, and videos that express their feelings. to create and put into practice a model that incorporates facial expressions as symptoms and indications. Based on that, it forecasts the reputation of the patient and whether they have been diagnosed with depression. We can train a model using pictures, and we'll utilize it to make predictions. After prediction, image captioning can be completed for better report visualization. We'll also employ text mining (NLP) technology to foretell melancholy using cues provided by people. We document the application of the suggested strategy. We used a variety of different psychological study facets to assess the efficacy of our suggested strategy. We demonstrate how the level of accuracy and error detection can be greatly increased by using our suggested approach. Using differential methods, tests are run on every combination of feature values listed in Table 1. (SVM with different characters, DT, and NB).

Any training or test sample should have the expected labels of depressed or not depressed. These conclusions were supported by tests that evaluated the relationship between characteristics and class labels. A variety of criteria are employed in combination to achieve the optimum separation outcomes. First, for training and testing, account measures with emotional elements (mixed, intermediate, and none of them) are used. The F1 phases of the different methods are summarized in Figure 4.

Figure 1 demonstrates that for all exploited classification methods, combining a mixed sentiment with the account measurements in their current state (as-is) results in a higher F1 measure. The F measure is 0.68 for DT and SVM-L and 0.784 for NB.

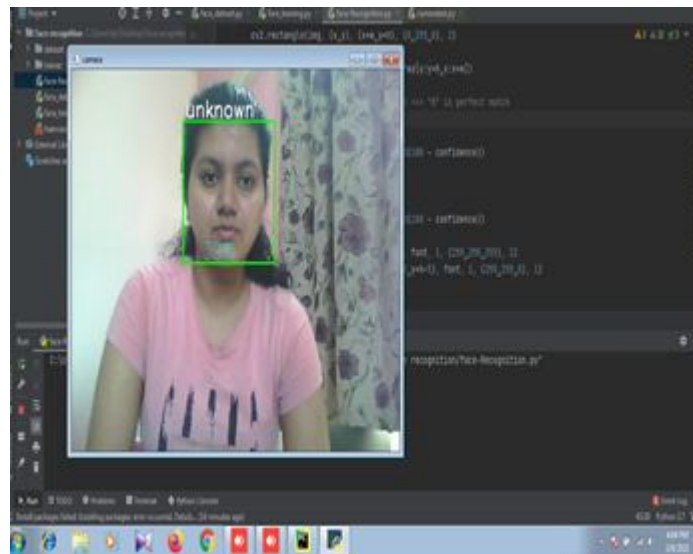


Figure 3: Face recognition

Menu
Home
Depression Detection
Testing

Depression Detection Using Machine Learning

Fill the Details

Enter Your First Name
Anil Kankar

Enter Your Email ID
anilkankar20@gmail.com

Enter your Contact No
789452361

How are you feeling today?

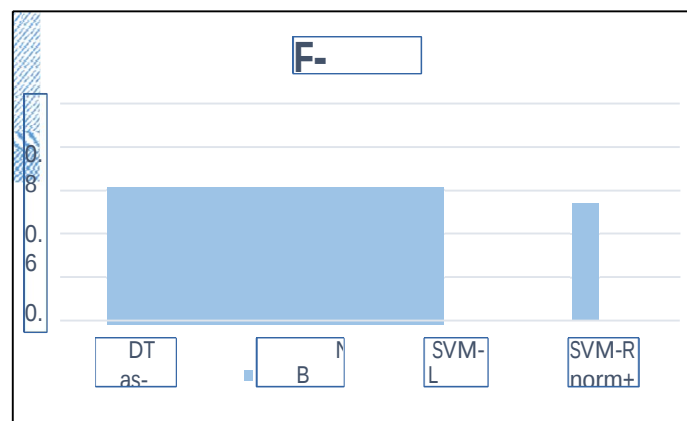
☐ excellent
☐ i am sad
☐ i am angry
☐ failure

Submit

Test Report

Depression Detected: 75.0%

Figure 4: Filling details of the user



VII. CONCLUSION

The predictor works reasonably well when used with data used to train the classifiers and for testing from the same dataset. The predictor frequently fails to pick up on the phrase linked with disdain, though this is mainly due to a lack of specific examples of disrespect in the training and test images, poor data labelling prior to training, and the difficulty in identifying disrespectful behavior. The classifier also fails to recognize emotions for test data that contain expressions that don't obviously correspond solely even when it hasn't been taught for

other expressions, to one of the seven basic phrases. Work in the future should concentrate on enhancing the classifiers' robustness by examining more precise detection techniques and utilizing more training photos from other datasets. taking into consideration the categorization of more sophisticated and nuanced phrases while also retaining computer efficiency.

REFERENCES

- [1] Social network analysis, Scott J. Sage, Thousand Oaks, 2017.
- [2] Social network analysis, Serrat O. Knowledge-based solutions. 39–43. Singapore: Springer; 2017.
- [3] Mikal J, Hurst S, and Conway M. [3]. A qualitative study exploring patient perceptions of the usage of social media data to support depression diagnosis and therapy. From linguistic signal to clinical reality: Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology. 2017.
- [4] social media, big data, and mental health: Recent advancements and ethical implications. Conway M, O'Connor D. 9: 77–82, Curr Opin Psychol, 2020.
- [5] Sentiment analysis in transcribed utterances. [5] Ofek N, et al. Data mining and knowledge discovery conference in the Pacific-Asia region. Springer, in Cham.
- [6] User interest and social impact-based emotion prediction for individuals. [6] Yang Y, et al. In: The 21st ACM International Conference on Multimedia Proceedings. 2019. ACM, New York.
- [7] The psychological meaning of words: LIWC and computerised text analysis approaches, Tausczik YR, Pennebaker JW. 2020;29(1):24–54 in J Lang Soc Psychol.
- [8] Pennebaker JW, Francis ME, Booth RJ. Linguistic inquiry and word count: LIWC 2020, vol. 71. Mahway: Lawrence Erlbaum Associates; 2001. p. 2020.
- [9] The early diagnosis of depression through social networking sites, Holleran SE, p. 9. 2020; Tucson: University of Arizona.
- [10] Emotion-focused therapy for depression. Per Centered Exp Psychother. 2021;16(1):106-17. Greenberg, L. S.
- [11] Prosperity and depression: a theoretical examination of cyclical motions, Haberler G. [11]. 2021; London: Routledge.
- [12] Guntuku SC, et al. An integrated review of depression and mental illness detection using social media. Curr Opin Behav Sci. 2020; 18:43–9.
- [13] Using social media to predict depression, De Choudhury M, et al. [13]. 2013. ICWSM, vol. 13, p. 1–10.
- [14] Predicting postpartum changes in emotion and behavior by social media. De Choudhury M, Counts S, Horvitz E. The SIGCHI conference on human aspects in computer systems proceedings. 2020; New York: ACM.
- [15] Detecting suicidality on Twitter, B. O'Dea et al. (2020). Internet Interv. 2(2):183–8. Signs Symptoms Behavior Refusing to go out anymore Not finishing tasks at work Not engaging in routine charming exercises unable to focus
- [16] Zhang L, et al. Using linguistic features to evaluate Chinese microblog users' likelihood of committing suicide. in: Human-centered computing international conference. 2019; Berlin: Springer.