

SVM Classification Technique to Analyze Mental Health and Stress Levels

Prof. P.M. Kamde¹, Shantanu Ramteke², Raturaj Kadam³, Gunj Hundiwala⁴, Suryansh⁵

Faculty, Department of Computer Engineering, Sinhgad College of Engineering, Vadgaon Bk. Pune¹

Students, Department of Computer Engineering, Sinhgad College of Engineering, Vadgaon Bk. Pune^{2,3,4,5}

Abstract: *Stress has become a serious problem in this current generation and the number of people affected by mental disorders is increasing day by day. However, some of them manage to acknowledge that they are facing depression while some of them do not know it. On the other hand, the vast progress of social media is becoming their “diary” to share their state of mind. Several kinds of research had been conducted to detect stress levels among various classes of people, but people normally shun this topic in shame and do not classify this problem as a disease to be treated. This Project based on analysis of mental health and aims to predict stress levels using visual inputs of the user, especially images and live camera and gives instant predictions about their mental state, i.e., happy, sad, disgusting, bored, worry, neutral, surprised. SVM model is used for feature extraction and classification into class labels. On visual inputs CNN is works best in comparison to other algorithms. Depression percentage is also shown giving exact measurements about a person's mental health.*

Keywords: Preprocessing, Feature Extraction, Segmentation, etc.

I. INTRODUCTION

Stress affects the everyday lifestyle in such a way that socially, academically, mentally, financially professionally disturbs an individual. When an individual is unable to balance between the strain that are placed on him/her and his/her ability to fight with them, then it causes pressure on psychological state which creates stress. Stress has become a serious problem in this current generation and the number of people affected by mental disorders is increasing day by day. Several kinds of research had been conducted to detect stress levels among various classes of people. Although psychologists prescribe various medications for their clients like anti-depressants, sleeping pills, etc.; still, the medication hasn't been able to cure or eradicate the sickness. There may be multiple reasons why a person is going through a certain situation like society, work pressure, family, etc. Our research on this topic will be limited to predicting such sickness in the human body and identifying what the person is going through using the previously recorded dataset. The infinite increase in population, the pressure of survival, and the pressure of learning makes the competition between people more and more fierce. Some college students have also been in a state of anxiety and panic for a long time, and mental health diseases have shown an explosive growth trend.

An alarming increase in suicide rate all over the world that has led to an undeniable need for depression detection. Early stages of depression can be cured by proper counselling. Emotion Recognition or Detection of human cognitive states is very important to address depression. Thus, a system which automatically recognizes emotional states of an individual using data from physiological signals certainly helps in mapping discrete human emotions to a certain pattern of physiological signals and data that can be analyzed for the cognitive state that it corresponds to. However, the task of accurate recording, effectiveness of the stimuli in eliciting target emotions and underlying emotions of the subjects that influence the perception of the target emotional state are certain challenges that should be overcome in implementing an emotion recognition system.

Users with mental illness have been identified by what they publicly share on online social media platforms and they can be distinguished from ordinary users by the language patterns in their written texts or other online activities. For example, depressed users tend to use more first-person singular pronouns in spoken language or depressed groups have elevated the use of the word “I” and negative emotions in comparison with control groups. So automated detection methods which are called early risk detection may help to detect depressed or in general mentally ill people through supervising of written text spreading in social medias, and then these individuals could be assessed more thoroughly to have been provided with resources, supports, and full treatments. Our automated detection machine is built by training a classifier using a set of available training data to categorize users into a control group and an affected group. This classification is performed using predictive models that use extracted features from users' written text or corpus.

II. MOTIVATION

Stress has been a new pandemic of this modern era where technology has advanced so much but mental health has degraded in the similar magnitude. People looking fit from outside are equally unfit from inside mentally. Suicidal tendencies have increased very much and life loss is significant. Machine Learning has done significant advancements in other fields then why not this mental health field. SVM algorithm's applications and utility has attracted us to use it into our project. Depression being a worldwide spread mental health ailment requires immediate and better ways to deal with. The emotion recognition algorithms have been implemented using physiological signals. Several techniques, including reduction in dimensionality of the feature set, retaining only the most distinctive features across emotions have been adopted to achieve better recognition accuracy of 91%. Although the physiological signal sources are less susceptible to individual-dependent variations, the recognition accuracy can certainly be increased if features from facial and speech signals are also added to the emotion recognition system.

Our work is based on social media data because it facilitates information sharing. We thought this to be helpful to analyse the data which the user is providing to us in order to make an early prediction of whether the person is suffering from depression or not. The industry has some other existing systems which are based on various other techniques like - Boosting with an accuracy of 75%, Support vector machines which give an accuracy of 85.71%, then there is some convolution neural network-based model with different kind of features fed into the model and this gives the accuracy of 95%, and another algorithm which is usually used is the random forest, which gave an accuracy of 81.04%. Different machine learning algorithms have been made use of for this purpose, which is as follows: Bayes theorem, Support vector machine, Logistic regression, and Random Forest. The accuracies achieved in ascending order are logistic regression with 96.2% accuracy, support vector machine with 96.1% accuracy, random forest with 95.1% accuracy and the Bayes theorem with the accuracy of 88.5% when we used the tf-idf technique and the accuracy of 83.15 when we used bag-of-words technique.

III. LITERATURE SURVEY

Efficient implementation of the nonlinearly separable data, the machine learning technique is used that is SVM. SVM has high-quality generalization for the binary classification problems. Finally, we conclude, the selection stage is crucial because the outcome rely on the kernel function; selection process the corresponding kernel function is a very difficult task, which is still an open problem up till now. Radial Basis Function is used for pattern recognition on linearly separable values. The experimental outcome on IRIS datasets we have discussed. Where we evaluate SVM based on the RBF method with other classifiers.

The report represents superiority to the presented approach (SVM based on RBF) which delivers extensively enhanced value against the other classifiers. The misclassification value is removed from the final Support Vector Classification. An improved accuracy is obtained after a certain modification on the epsilon and cost parameter in radial basis function.

It has been observed in the experiment that the trained SVM gives 98% more accurate results as compared to the kernel tricks. However, for predicting and training the labels of the class this proposed method takes longer time which serves as its shortcoming. SVM classifications are finding more and more uses in different fields like image recognitions. [1]

Our research on this topic will be limited to predicting the mental sickness in the human body and identifying what the person is going through using the previously recorded dataset. We will be using Logistic Regression, Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbor, and Naïve-Bayes algorithms for creating ensemble models and further compare the models. We have applied the proposed algorithms on the Kaggle dataset having 334 sample sizes with 31 different fields about unemployment and mental illness. In the end, the test result of this application can be an authentic example of IoT in healthcare. [2]

This study uses text-level mining of Sina Weibo data from college students to detect depression among college students. First, collect text information of college student users in Sina Weibo, and construct the text information into input data that can be used for machine learning. Deep neural networks are used for feature extraction. A deep integrated support vector machine (DISVM) algorithm is introduced to classify the input data, and finally realize the recognition of depression. DISVM makes the recognition model more stable and improves the accuracy of depression diagnosis to a certain extent. Simulation experiments verify that the proposed depression recognition scheme can detect potential depression patients in the college student population through Sina Weibo data. [3]

In our work we selected ECG as the bio signal and extracted its features. The advantage of taking ECG as the bio signal is, information about respiratory signals - EDR (ECG Derived Respiration) feature can be easily derived without any extra sensors. Among those unique features we chose ECG derived Respiration, Respiration Rate, QT interval. For training and validation of our new model we used Physionet's "drivedb" database. Our proposed model uses Optimised Support Vector Machines (SVM) using decision trees. Our experimentation results show better accuracy in detecting stress. [4]

In this study, we implemented a single-lead ECG system based on reactivity of HRV, and combined mental task and paced deep breathing thereby improving the screening accuracy. Moreover, support vector machine (SVM) model was applied for classification of HRV indices. We tested the system on 16 healthy subjects and 7 psychiatric patients with depression or somatoform disorder. A significant difference was found between the healthy group and the patient group for the response of the HRV indices on dual mental tasks. The SVM non-linear classification model achieved a sensitivity of 71.4% and specificity of 93.8%. [5]

Various social networking platforms are used to allocate various data and routines with others. This can prove to be really helpful in understanding the mind of a depressed person. Machine learning algorithms have proved to be quite helpful in the past where researchers have worked on social media data to predict the number of people suffering from depression based on their early symptoms and social media activity. The Method of objective is to divide the approach into two parts - the first one is based on the time and writing patterns of the content and the second one is based on the linguistic clues, analysing the text or the tweet which has been shared. The objective is to help patients suffering from this disease in the early detection of the symptoms of depression which could prove beneficial to them and to their family too. [6]

Support Vector Machine (SVM) has been used for classification because of its high recognition rates and superior performance compared to Bayesian and Regression-based classifiers. The data corresponding to eight emotions available in databases DEAP, MAHNOB-HCI has been used for training and testing the proposed system. The physiological signals namely Electromyogram (EMG), Blood Volume Pressure (BVP) and Galvanic Skin Response (GSR) from emotion Sentic database are considered. Classification accuracy of 75% has been obtained for five target emotions, namely, Joy, Grief, Anger, Hate and Reverence. An improved recognition rate of 91% has been obtained by using k-fold leave out one cross-validation to reduce the observation dependence of prediction when there is limited training data. [7]

In this paper, we study mental illnesses through people's choice of words in expressing themselves on two popular social media platforms, Reddit and Twitter. Our goal is to develop an empirical model to detect and diagnose major mental disorders in individuals. We build a substantial dataset of posts made by people suffering from mental illnesses and the control ones, and in order to generate numerical feature from text we apply text cleaning and Word2Vec language modelling, and then for classification we used SVM machine which classifies posts and users with high accuracy. We achieve an accuracy of 95% on Twitter users and an accuracy of 73% on the Reddit challenge. [8]

A fair number of studies have been made and their results are published on how social network data can help and improve the prediction of sentiment analysis. This study sets out to assess whether the text analyzing of online user posts such as Twitter or Reddit can be of use for finding users who suffer from mental illnesses. The main obstacle we had to overcome was finding a way to conveniently extract useful information from such a large number of messages for each user. To this purpose, we have applied Word2Vec model and some feature selection techniques, namely Fisher Ratio and multi-level classifiers that allow us to tell mentally ill users from the control ones with great precision. [9]

This paper defines a binary classification problem as identifying whether a person is depressed, based on his tweets and Twitter profile activity. Different machine learning algorithms are exploited and different feature datasets are explored. Many preprocessing steps are performed, including data preparation and aligning, data labeling, and feature extraction and selection. The SVM model has achieved optimal accuracy metric combinations; it converts an extremely nonlinear classification problem into a linearly separable problem. Although the DT model is comprehensive and follows understandable steps, it can fail if exposed to brand-new data. This study can be considered as a step toward building a complete social media-based platform for analyzing and predicting mental and psychological issues and recommending solutions for these users. The main contribution of this study lies in exploiting a rich, diverse, and discriminating feature set that contains both tweet text and behavioral trends of different users. This study can be extended in the future by considering more ML models that are highly unlikely to over-fit the used 15 data and find a more dependable way to measure the features' impact. [10]

The ML approaches can be used to assist in diagnosing mental health conditions. PTSD, schizophrenia, depression, ASD, and bipolar diseases lie in the domains of mental disorders. Social media data, clinical health records, and mobile devices sensors data can be analyzed to identify mood disorders. In this paper, we surveyed state-of-the-art research studies on the diagnosis of depression using ML-based approaches. The purpose of this review paper is to provide information about basic concepts of ML algorithms frequently used in the mental health domain, specifically for depression and their practical application. Among the reviewed studies, SVM has been the most used classifier for detecting depression as it works well with unstructured and high-dimensional data. SVM is also resistant to overfitting. SVM can be proved to be an efficient algorithm for data with an anonymous and irregular distribution. As anticipated, most of the SVM classifiers developed in the articles had a high accuracy of greater than 75%. Because data in the mental health area are scarce, SVM outperforms other machine learning methods for diagnosis. We discussed some of the MHMS's research difficulties and potential advancements in mental health and depression. According to the research reviewed, applications based on machine learning provide a significant potential for progress in mental healthcare, including the prediction of outcomes and therapies for mental illnesses and depression. [11]

This depression detection model can help detect people suffering from depression. According to the literature survey, there are separate BERT and SVM models for depression prediction, which achieved a fair accuracy. Hence to enhance the accuracy more, in the proposed work, an ensemble model (BERT and SVM) was developed. With the user giving the inputs, a decision is made whether the user is under depression or not. Thus, the model for depression prediction successfully categorizes depressed and non-depressed people. [12]

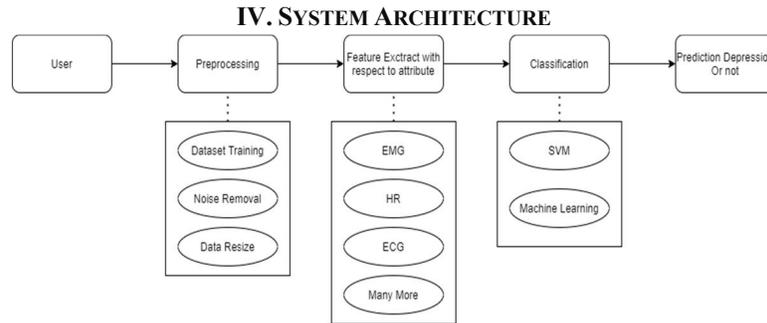
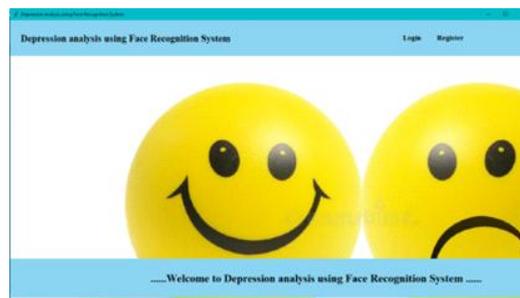


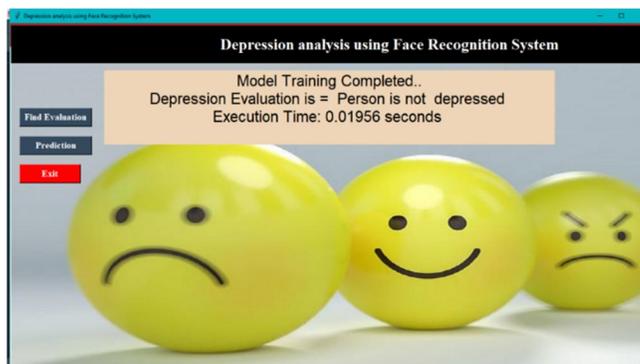
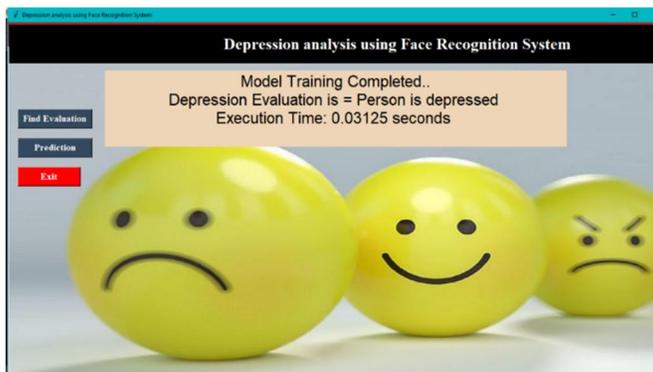
Figure: System Architecture

V. ALGORITHM

SVM: In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data for classification and regression analysis. Support vector machines (SVMs) are a set of supervised learning methods used for classification, regression and outliers' detection. The advantages of support vector machines are: Effective in high dimensional spaces. Still effective in cases where number of dimensions is greater than the number of samples. Support Vector Machine (SVM) is a supervised machine learning algorithm capable of performing classification, regression and even outlier detection. The linear SVM classifier works by drawing a straight line between two classes. A kernel is a function used in SVM for helping to solve problems. They provide shortcuts to avoid complex calculations. The amazing thing about kernel is that we can go to higher dimensions and perform smooth calculations with the help of it. We can go up to an infinite number of dimensions using kernels.

VI. RESULTS





VII. CONCLUSION

Successful classification of image dataset provided by the user is done and the user is given the label according to the parameters detected by the SVM algorithm and the output is displayed successfully. Categories like Happy, Sad, Fear, Surprised, etc. is displayed in the output. We have observed that SVM algorithm provides better accuracy in image analysis than other algorithms.

REFERENCES

- [1] Support Vector Machine Accuracy Improvement with Classification. Lalit Mohan Gehu, Bhimtal, Uttarakhand, Janmejy Pant Gehu, Bhimtal, Uttarakhand, Priyanka Suyal MIET Haldwani, Uttarakhand, Arvind Kumar Uttarakhand Technical University, Dehradun (12th International Conference on Computational Intelligence and Communication Networks).
- [2] A Machine Learning Implementation for Mental Health Care, Application: Smart Watch for Depression Detection: Piyush Kumar com, Dr. Achyut Shankar, Rishi Chauhan, Dr. Sanjeev Thakur, Dr. Thompson Stephan. (2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence 2021)).
- [3] A Depression Recognition Method for College Students Using Deep Integrated Support Vector Algorithm. Corresponding author: Xuemei Chen (Received March 31, 2020, accepted April 9, 2020, date of publication April 13, 2020, date of current version May 5, 2020.).
- [4] A Decision Tree Optimised SVM Model for Stress Detection using Bio signals, Alana Paul Cruz, Aravind Pradeep, Kavali Riya Sivasankar and Krishnaveni K.S, International Conference on Communication and Signal Processing, July 28 - 30, 2020, India.
- [5] Development of a Mental Disorder Screening System Using Support Vector Machine for Classification of Heart Rate Variability Measured from Single-lead Electrocardiography Authors: Mai Kobayashi, Takemi Matsui, Guanghao Sun, Tetsuo Kirimoto, Toshikazu Shinba.
- [6] EARLY DETECTION OF DEPRESSION FROM SOCIAL MEDIA DATA USING MACHINE LEARNING ALGORITHMS G. Geetha, G. Saranya, Dr. K. Chakrapani, Dr. J. Godwin Ponsam, M. Safa, Dr. S. Karpagaselvi (2020 2nd International Conference on Power, Energy, Control and Transmission Systems).
- [7] Detection of Mental Illness Risk on Social Media through Multi-level SVMs Kimia Hemmatirad , Hojjat Bagherzadeh, Ehsan Fazl-Ersi, Abedin Vahedian. (2020 8th Iranian Joint Congress on Fuzzy and intelligent Systems (CFIS), September 2-4, 2020, Mashhad, Iran).
- [8] Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., Chua, T.-S., & Zhu, W. Depression detection via harvesting social media: A multimodal dictionary learning solution in IJCAI 3838-3844 (2017).
- [9] Vitriol, V.; Cancino, A.; Weil, K.; Salgado, C.; Asenjo, M.A.; Potthoff, S. Depression and psychological trauma: An overview integrating current research and specific evidence of studies in the treatment of depression in public mental health services in chile. *Depress. Res. Treat.* 2014, 2014, 608671.
- [10] Arora, Priyanka, and Parul Arora. "Mining twitter data for depression detection." 2019 International Conference on Signal Processing and Communication (ICSC). IEEE, 2019.