Intelligent Depression Detection and Support System that Includes Sentiment Analysis from Social Cues using Aspect Mining and Deep Learning

Gunjali Gajare[1], Rajeshri Tarde[2], Vaishnavi Borate[3], Pooja Bramhane[4], Prof. Sachin Korde[5]

[1,2,3,4] Students & [5]Asst. Prof. of Department of Information Technology,

Savitribai Phule Pune University, Pune, Maharashtra, India

Abstract: Online social networks provide relevant information on users' opinions and posts on various topics. So Applications, like monitoring and detection systems can collect and analyze this data. This paper study an information-based system, which incorporates an emotional health monitoring system to detect users with possible psychological disorders especially depression and stress. Symptoms of this mental disorder are usually observed passively. During this situation, the author argues that online social behavior extraction offers a chance to actively identify psychological disorders at an early stage. It's hard to acknowledge the confusion in light of the very fact that the mental components considered within the standard symptomatic criteria survey cannot be seen by the registers of online social exercises. Our methodology, new and artistic for the act of psychological disorders. Location does so don't open up to the self-revelation of these mental factors through the surveys. Rather, propose an AI strategy that's recognition of psychological disorders. In informal organizations which misuses the highlights far away from interpersonal organization information for concerning exactness potential instances of confusion discovery. We perform an analysis of the characteristics and that we also apply machine learning in large-scale data sets and analyze features of the kinds of psychological disorders.

Index Terms - Sentiment analysis, Knowledge Personalization and Customization, Detection System, Social Networks, Machine Learning, Social Cues, Deep Learning, Aspect Mining, etc

I. INTRODUCTION

In many situations, humans who are depressed are ignorant of their disturbed mental condition. They are unable to identify the cause of constant unhappiness in them and eventually such students fall into a state of mind where they start having suicidal tendencies. In some cases, students do know that they are suffering from depression, but they are hesitant to seek any kind of help from anyone mainly due to the wrongly conceived notion of 'humiliation' associated with depression. It is better to identify the signs of depression at the initial stages of depression. Depression if identified in the initial stages, just a simple one-hour talk with a counselor may be of immense help

for the student. This may change the negative state of mind of that student into a positive one. Such a student can be given good counseling on how to deal with mental stress and can be guided to follow the right path to success. The most important form of non-verbal communication is the facial expressions of a person. Many studies have been done for finding out the facial expressions that are related to depression. The current work is mainly undertaken to find out the presence of depression in college students by studying their facial features. This system mainly uses different image processing techniques for face detection, feature extraction, and classification of these features as depressed or non-depressed. The system will be trained with features of depression. Then videos of different students with a frontal face will be captured using a web camera. Then the facial features of these faces will be extracted for prediction of depression. Based on the level of depression features the student will be classified as depressed or non-depressed.

Feature selection methods should be classified into different aspect categories like wrapper, static and hybrid techniques. In filter-based approaches, the selection of features can't reliant on any machine learning algorithm. In this, features are preferred based on their numerical weight. In the dynamic approach, first different subsets of features are identified then are evaluated using one of the classifiers. The hybrid approach is the combination of different feature extraction as well features selection methods, it also used various machine learning algorithms, In-univariate filter approach features are evaluated concerning relevance. The multivariate method considers the correlation between features and avoids redundant features. We are proposing a filter approach that selects relevant features and avoids redundant ones.

II. RELATED WORK

- In certain studies, patients were given wearable devises to monitor their physical health, emotional behaviour and social interaction for identifying depression [6]. Some researchers have collected datasets by showing individuals film-strips to capture the facial expressions of subjects watching them. Data is also collected by giving a task of recognizing negative and positive emotions from different facial images [7]. Rather than analyzing a video for depression detection frame by frame, better results have been got for detection of depression when the video is considered as a whole. [8] For this the patient's face region is first initialized manually. Then KLT (KanadeTomasi-Lucas) tracker is used to track the face throughout the video. The KLT tracker extracts curvature information from an image, i.e. for a sad expression the corners of the mouth would be angled down. Video based approach showed more accuracy as it generalizes the face region more accurately and so the minute movements within the face region are also considered for depression detection.

- Many studies have been conducted to identify the precise facial expressions that are related to depression. A study has been conducted for finding out Action Units (AU) related to different emotions exhibited by depressed patients [1]. The presence of AU12 which is associated with emotion smile was low in highly depressed patients. The presence of AU14 related to emotion contempt and AU10 related to emotion disgust was also present along with AU12. The video data for this study was collected through clinical interviews of depressed patients as well as non-depressed patients. The results showed that AU14 related to emotion contempt proved most accurate for depression detection

- Another study presented a technique which uses analysis of facial geometry along with analysis of speech for depression detection [3]. This work says that the expressions associated with depression are found to be in

lower frequencies in smaller duration videos. Therefore, longer time videos need to be captured for effective depression detection. Datasets are also created by capturing videos of patients while answering clinical interviews. Interviews recorded were for both for depressed patients as well as non-depressed patients. Videos are also recorded from the diagnosis of depression till the patient has improved. [1][4]. Studies showed that there is a significant relation between facial features and vocal behavior of the depressed [5].

- The students suffering from depression would show less attentiveness in classrooms. If the students' emotions are mapped to the activities done in classroom, their emotional state can be found out whether they are depressed or not, and based on this the teacher can help the student by paying more attention to that particular student. [11] If different faces in the same scene show the same positive or negative sentiment, it would help to understand the whole situation of the scene, whether subjects in the scene are happy or whether something wrong is happening in the scene [12].

- Features related to eye movement to understand the eye activity of the depressed and features related to head pose movement to understand the head movement behavior of the depressed has been done in [2]. The classification of the features related to eye activity showed higher significance in detecting severe depression. Detection of depression from facial features can be done by measuring 'Multi-Scale Entropy' (MSE) on the patient interview video. [4] MSE helps to find out the variations that occur across a single pixel in the video. The entropy levels of highly expressive, non-depressed patients were high. The entropy level was low for depressed patients who were less expressive of their emotions.
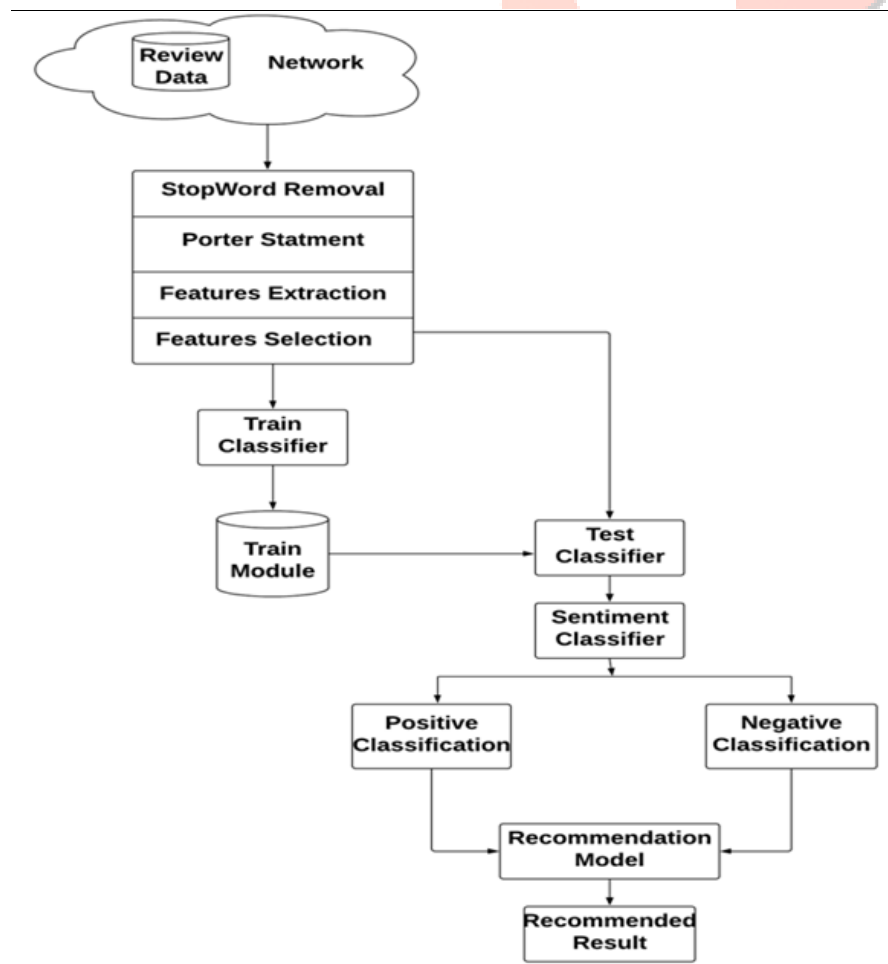
III. PROPOSED SYSTEM

Fig.1: Proposed System Architecture

In the proposed systemic approach, we formulate the task as a classification problem to detect two types of detection of psychological disorders in social networks using the machine learning framework: i.e. first is **Stress** and second is **Depression.**

An innovative solution to monitor and detect potential users with emotional disorders, according to the classification of sentences with depressed or stressed content. Machine learning is used for representation at the character level and the recognition of the extent of the disturbance.

Sentiment analysis is a huge term to classify user's opinion using Natural Language Processing (NLP) and Machine Learning (ML) Approach. Various researchers have done different methods for positive negative classification, aspect base classification, polarity based classification etc. Product review based sentiment analysis is similar to proposed sentiment analysis approach.

IV. RESEARCH METHODOLOGY

The above fig.1 illustrates the system architecture; it carried the overall classification process on sentiment analysis using movie review data set. Initially pore-processing analyzes the opinions from syntactical point of view and original syntax of sentence is not disturbed. In this phase, the several techniques like POS tagging, Stemming and Stop word removal are applied to data set for noise reduction and facilitating feature extraction.

1. **Data Pre-processing**

   In the data pre-processing phase, we first process the data which is extracted from training as well as testing documents. Various methods have been used for data pre-processing these are describe in below section

2. **Stop Word Removal:**

   Stop words are common and high frequency words like "a", "the", "of", "and", "an". Different methods available for stop-word elimination; ultimately enhance performance of feature extraction algorithm.

3. **Stemming**

   Stemming and Lemmatization are two essential morphological processes of pre- processing module during feature extraction. The stemming process converts all the inflected words present in the text into a root form called a stem. For example, 'automatic,' 'automate,' and 'automation' are each converted into the stem 'automat.' Stemming gives faster performance in applications where accuracy is not major issue.

4. **Lemmatization (lemmas):**

   The lemma of a word includes its base form plus inflected forms. For example, the words "plays", "played and "playing" have "play" as their lemma. Lemmatization groups together various inflected forms of word into a single one. Stemming also removes word inflections only whereas; Lemmatization replaces words with their base form. For example, the words "caring" and "cars" are reduced to "car" in a stemming process whereas lemmatization reduces it to "care" and "car" respectively, hence lemmatization is considered to be more accurate.

5. **Part of speech (POS) tagging:**

   Parts of speech or POS tagging is a linguistic technique used which is used many existing researchers, for product feature extraction as product aspects are generally nouns or noun phrases. POS tagging assigns a tag to each word in a text and classifies a word to a specific morphological category such as norm, verb, adjective,

etc. POS taggers are efficient for explicit feature extraction in terms of accuracy they achieved, however problem arises when review contains implicit features.

6. **Features Extraction**

In this phase system extract various feature set using machine learning methods for sentiment classification. We extract four basic features from preprocessed data like unigram features, Bi-tagged features, dependency rule base features etc. all these feature extraction techniques have illustrated in below section.

7. **Unigram features**

A Unigram feature is nothing but a bag of words extort by separating text by spaces as well as noise characters. The unigram model is additionally called the bag of words model.

8. **Features Selection**

Hybrid method has used for feature selection from full extracted features. Basically three types of features have been extracted from given data. The purpose of select the best feature which increase the accuracy of classification. Many irrelevant features appear during the feature extraction; it need to eliminate when we select the features. We used TF-IDF, Maximum Relevance and co-relation base hybrid method has used to select the features. The benefit of this method provides respective features selection for individual features set. The TF-IDF cosine similarity, TF-IDF Co-occurrence matrix and MRMR method has used for feature selection.

9. **Classification:**

After we get the training model, we can feed the testing data into it and get the prediction of classification. The testing stage includes preprocessing of testing text, vectorization and classification of the testing text.

IV. RESULTS AND DISCUSSION

Results and Analysis

The system can provide the below outcomes once successfully executed training as well as testing phase

- System gets background knowledge as a training set
- System can able to find weight vector and define the similarity with respective domain.
- Reduce false positive ratio.
- Each cluster has categorized into multiple similar clusters, under the one master cluster.
- Finally, similarity score will classify each bucket into the respective domain.

V. CONCLUSION

In this paper, automatically identifying potential online users with depression and stress is threatening people's health. It is not trivial to detect the disorder in time for proactive care. Therefore, we present a framework to detect the states of stress and depression of users from the monthly data of users' social networks, exploiting the content of Facebook publications and the social interactions of users. Using the data of the social networks of the real world as a basis, we study the correlation between the states of psychological disorder of users and their social interaction behavior; we recommend the user for health precautions to send by mail for user interaction.

ACKNOWLEDGMENT

REFERENCES

[1] Girard, Jeffrey M., Jeffrey F. Cohn, Mohammad H. Mahoor, Seyed mohammad Mavadati, and Dean P. Rosenwald. "Social risk and depression: Evidence from manual and automatic facial expression analysis." In Automatic Face and Gesture Recognition (FG), 10th IEEE International Conference and Workshops on, pp. 1-8. IEEE, 2013.

[2] Alghowinem, Sharifa, Roland Goecke, Jeffrey F. Cohn, Michael Wagner, Gordon Parker, and Michael Breakspear. "Cross-cultural detection of depression from nonverbal behaviour." In Automatic Face and Gesture Recognition (FG), 11th IEEE International Conference and Workshops on, vol. 1, pp. 1-8. IEEE, 2015.

[3] Pampouchidou, A., O. Simantiraki, C-M. Vazakopoulou, C. Chatzaki, M. Pediaditis, A. Maridaki, K. Marias et al. "Facial geometry and speech analysis for depression detection." In Engineering in Medicine and Biology Society (EMBC), 39th Annual International Conference of the IEEE, pp. 1433-1436. IEEE, 2017.

[4] Harati, Sahar, Andrea Crowell, Helen Mayberg, Jun Kong, and Shamim Nemati. "Discriminating clinical phases of recovery from major depressive disorder using the dynamics of facial expression." In Engineering in Medicine and Biology Society (EMBC), 38th Annual International Conference of the, pp. 2254- 2257. IEEE, 2016.

[5] Cohn, Jeffrey F., Tomas Simon Kruez, Iain Matthews, Ying Yang, Minh Hoai Nguyen, Margara Tejera Padilla, Feng Zhou, and Fernando De la Torre. "Detecting depression from facial actions and vocal prosody." In Affective Computing and Intelligent Interaction and Workshops. ACII 2009. 3rd International Conference on, pp. 1-7. IEEE, 2009.

[6] Tasnim, Mashrura, RifatShahriyar, NowshinNahar, and Hossain Mahmud. "Intelligent depression detection and support system: Statistical analysis, psychological review and design implication." In e-Health Networking, Applications and Services (Healthcom), 18th International Conference on, pp. 1-6. IEEE, 2016.

[7] Pampouchidou, Anastasia, Kostas Marias, ManolisTsiknakis, P. Simos, Fan Yang, and FabriceMeriaudeau. "Designing a framework for assisting depression severity assessment from facial image analysis." In Signal and Image Processing Applications (ICSIPA), International Conference on, pp. 578-583. IEEE, 2015.

[8] Maddage, Namunu C., RajindaSenaratne, Lu-Shih Alex Low, Margaret Lech, and Nicholas Allen. "Video-based detection of the clinical depression in adolescents." In Engineering in Medicine and Biology Society, (EMBC). Annual International Conference of the IEEE, pp. 3723-3726. IEEE, 2009.

[9] Meng, Hongying, Di Huang, Heng Wang, Hongyu Yang, Mohammed AI-Shuraifi, and Yunhong Wang. "Depression recognition based on dynamic facial and vocal expression features using partial least square regression." In Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge, pp. 21-30. ACM, 2013.

[10] Katikalapudi, Raghavendra, Sriram Chellappan, Frances Montgomery, Donald Wunsch, and Karl Lutzen. "Associating internet usage with depressive behavior among college students." IEEE Technology and Society Magazine 31, no. 4 : 73- 80. 2012

[11] Sahla, K. S., and T. Senthil Kumar. "Classroom Teaching Assessment Based on Student Emotions." In The International Symposium on Intelligent Systems Technologies and Applications, pp. 475-486. Springer International Publishing, 2016.

[12] Athira, S., R. Manjusha, and Latha Parameswaran. "Scene Understanding in Images." In The International Symposium on Intelligent Systems Technologies and Applications, pp. 261-271. Springer International Publishing, 2016.