International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Prediction of Bipolar Disorder with Voice Analysis using Machine Learning Techniques

K. Naga Yaswanth Reddy, Ganesh Naga Sai Nerella, K. Samantha Roy, Swarna Kuchibhotla, B. Manjula Josephine

Abstract*: The change in the speech is the responsive and well-founded measure of the depression and obsession of the bipolar disorder. This analysis mainly focuses on perceiving the voice attributes and phone calls data is collected as it acts as a main search-space maker for bipolar clutters. By combining the voice features with the phone call data based on their behavioral activities, self- monitoring data control and illness activities the accuracy would increase to effective states. The voice attributes and smartphones collect the activities of sample phone data and self-monitoring data. These activities are the root cause of the expansion of two symptoms: depression and obsession. These symptoms were introduced by a researcher who was rendered with smartphones. The phone call data were examined through a statistical random forest algorithm. The states were extracted from daily phone calls and are classified using voice attributes. These attributes are more determined and accurate to classify the maniac states. The main subject in comparing the voice attributes and self-observed data with the behavioral activities of phone call data is that these attributes increase the efficiency, vulnerability, and definiteness of classifying the affective states. The techniques used to detect the voice features are support vector machine (SVM) random forest. the proposed system will enhance the performance of the prediction of all the techniques. By comparing all these techniques by finding the accuracy of each technique we can know which technique predicts more accurately.

Keywords: Bipolar disorder, voice analysis, Classification, Voice datasets.

Manuscript published on November 30, 2019.
* Correspondence Author
K. Samantha Roy*, Dept of Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India.
Ganesh Naga Sai Nerella, Dept of Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India
K. Naga Yaswanth Reddy Dept of Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India.
Swarna Kuchibhotla Assoc Professor, Dept of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation,Guntur, India
B.Manjula Josephine, Assitant Professor, Dept of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur , India.

LITERATURE SURVEY:

Table 1: Literature Survey of various techniques used in the prediction of bipolar disorder.

Year	Publisher	Objective
1967	Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol	This is an account of further work on a rating scale for depressive states, including a detailed discussion on the general problems of comparing successive samples from a 'population'
1978	Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity, and sensitivity. *Br J Psychiatry*	An eleven-item clinician-administered Mania Rating Scale (MRS) is introduced, and its reliability, validity, and sensitivity are examined. There was a high correlation between the scores of two independent clinicians on both the total score (0.93)
1938	Newman S, Mather VG. Analysis of the spoken language of patients with affective disorders. *Am J Psychiatry*	Clinical descriptions and language analyses are presented of four patients with affective disorders.
2012	Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it?	Patients with mental disorders show many biological abnormalities that distinguish them from normal volunteers; however, few of these have led to tests with clinical utility.

I. INTRODUCTION

The audio-visual emotion challenge is the competition aimed at comparing the processed audio data and different classification algorithms for automatic voice, video analysis and speech health, and emotion sensing. The main aim is to bring together multiple communities from different disciplines, in particular the multimedia communities and those in the psychological and social sciences study expressive behavior .another objective is to provide a common benchmark test set for multimodal information processing,

to compare the relative merits of the approaches to automatic health and emotion analysis under well-defined conditions with large volumes of un-segmented, non-prototypical and non-preselected data of fully naturalistic behavior. For this project, a prior important concept is voice analysis.

The overall concept of the project is to identify the people who are suffering from bipolar disorder for this we first need some data. This data can be in the form of audio, video, etc. This data is collected from particular datasets like RECOLA, SWEA, etc. some finite number of features are extracted from these datasets and these extracted features are classified using some classification and regression algorithms used in machine learning. here in this, we are mainly focusing on support vector machine (SVM), random forest by which we are comparing these algorithms on the single dataset and proving one technique is best This classification shows the average amount of people suffering from Bipolar Disorder[1-3]. Literature Survey of different techniques used in prediction of bipolar disorder is given in the table 1[3-5].

A. Bipolar Disorder:

Bipolar disorder is a mental health illness. It is mainly due to the high maniac or depression of the individual. There is no proper medicine for bipolar disorder. Bipolar features are angry, depressed, very happy, low energy and unusual mood swings in the voice. There is no proper medicine for treating people with bipolar disorder. People with bipolar disorder can be examined thoroughly by interviewing patients weekly or daily basis. People with bipolar disorder have a neurological disorder at some point in their lives, It is a lifelong disease with high suicidal risk. The World health organization has ranked bipolar disorder as the top 10 diseases for adults

B. Bipolar Disorder corpus:

The corpus callosum abnormalities in patients with this bipolar disease are scanned through an MRI scan. There are huge differences between the regional corpus areas of the patients. The corpus is the hemispheric white tissue pathway that consists of around 200 million axons. This corpus develops in the children and slowly rises to adulthood. The results of the patients become more heterogeneous who are having this disorder. These results may lead to the variability in the phenotypic proportion of the patients. The video recordings are given in their entirety, without task-based segmentation, but atone is played when the task is witched, produced by the system that presented the instructions on the screen. The audio and video of the speech analysis is represented in form of dataset.

C. Symptoms and signs of bipolar disorder:

There may be many symptoms of the disease. When present, common symptoms of bipolar disorder may include:

- Extended periods of feeling overly happy or outgoing
- Extremely irritable mood
- Agitation, or jumpiness
- Talking very fast
- Jumping from one idea to another
- Being easily distracted
- Increasing goal-directed activities
- Restlessness.

II. DATABASES USED

A. RECOLA (The Remote Collaborative and Affective Interactions)

The RECOLA database is the very first of its kind and might be of great interest to all research groups working on the automatic sensing of social and affective behaviors expressed by humans in real-life conditions from multimodal cues. Recola is used to recorded to study socio-affective behaviors from multimodal data. Recordings, features, annotations, a temporal arrangement of events (face or speech detection) and data are provided for the twenty-three participants from the coaching and development partitions in numerous modules. There are twelve modules in total within the information repository, split over five main folders: Annotation (1), Audio (4), Bio-signals (2), data (1) and Video (4). An outline of the content of those modules is given below; the reader has spoken the subsequent article for any details on the computation of features: Pattern Recognition Letters, 2015. additionally, to other sets of options, and developed for the AVEC'16 and AVEC'18 Challenges are accessible within the information repository. Details relating to the info assortment and annotation are given within the article introducing the RECOLA dataset and bestowed at Face & Gestures 2013. Details relating to the annotation of the laughter events are given within the paper bestowed at ICMI'18[5-8].

B. SEWA

SEWA is a database that consists of different recorded voice data and video data to represent dynamic behavior that is recorded by standard webcams and the collected data of audio and videos are represented as SEWA projects.

These data are collected by volunteers that have been divided into groups based on their interests, age, gender and research groups. The experiment is divided into two groups and all volunteers are asked to participate in it and result in two sets of recordings.

Experiment Part 1: A volunteer is asked to observe a minimum range of four patients. These adverts are chosen to elicit mental states together with amusement, empathy, liking, and tedium. once observance of the advert, the volunteer is additionally asked stand-in a form to self-report his/her emotion and sentiment toward the advert.

Experiment Part 2: once observance the fourth one, he/she discusses the last watched advert and also the content he/she has simply seen with another volunteer typically best-known to the primary volunteer through a video-chat code.

The discussion is meant to elicit any reactions and opinions concerning the advert and also the publicized product like whether or not the merchandise is to be purchased, whether or not it's to be suggested to others, what are the most effective elements of the advert, whether or not the advert is suitable, however it are often increased, etc.. once the discussion, every participant is asked to fill during a form to self-report his/her emotion and sentiment toward the discussion.

The entire observance of patients and also the later speech between the volunteers are recorded victimization web-cameras and microphones integrated into the laptops/PCs of the volunteers[8-10][16].

III. FEATURES

A. MFCC

MFCC means Mel-Frequency Cepstral constant. It is most significant and widely used technique used in MATLAB to extract the features from audio or speech signals that is passes as a .wav file type. The speech signal that passes as input is divides into frames using audio buffers or an hamming window. while, applying the hamming window set frame overlap to zero as the voice data using in classification is noise free. The detail feature extraction process is as follows.
1) Framing: It is the method used to convert audio signals to frames using audio-read function in MATLAB which divides the audio into frames.
2) Windowing: It is used to minimize the signal discontinuity of frames (achieved using framing) before beginning and after finish Hamming window good choice for windowing in MFCC.
3) Fast Fourier transformation (FFT): FFT is used to covert the audio signal that is divided into frames to represent frequency and vice versa. It calculates rapidly by factoring the audio matrix with low complexity.
4) Mel scale filter bank: It is used to capture energy of a audio signal frame. It provides better results at low frequency.
5) Log energy computation: It is used to provide better audio and noise free signals for processing the data. Filter bank is applied to energy results in high spectral frames.
6) Cepstral means correction: Cepstral mean and variance normalization is a computational technique for audio processing and it differentiates the mean of buffer frames.
7) Discrete Cosine Transformation: DCT is used to express the data points of audio signals as real world values in a matrix.
8) The audio signal is finally processed to produce the noise free values using MFCC[14-15].

B. Pitch

Pitch is that the rise and fall of our voice once we speak, generally known as "highness" or "lowness" and that we use to pitch to offers delicate intending to sentences and that we pitch is one of the options accustomed extract the emotional disturbance

C. GTCC

The abbreviation of GTCC is Gammatone cepstral coefficients. That are biologically galvanized options computed by applying gammatone filter bank to the buffer of the signal, followed by the appliance of the exponent and also the separate trigonometric function transformation. which are antecedent utilized within the field of speech analysis, are custom-made for non-speech audio such as whistling, humming or hissing the data on classification function. Their performance is evaluated on to audio corpora of four h every (general sounds and audio scenes), following 2cross-validation schemes and 4 machine learning strategies in step with the results, classification accuracies are considerably higher once using GTCC instead of alternative progressive audio options [8-11]. As an in-depth analysis shows, with an identical procedure price, the GTCC is simpler FCC in representing the buffer characteristics of non-speech audio signals, particularly at very low frequencies. Intensity is that the measurable quantity of a property and it's one quite of feature extraction employed in manic depressive illness the intensity worth is recorded and with the assistance of the intensity worth we have tendency to calculate whether or not the person is stricken by manic depressive illness Grayscale pictures A component (short for image element) could be a little block that represents the number of grey intensity to be displayed for that individual portion of the image for many pictures, components values are integers that varies from 0(black) to 255 (white). The vary of intensity values from 0(black) to 255 (white). The computed intensity values represent the densities of the scanned object.[12-13]

IV. COMPARITIVE ANALYSIS

In this literature survey on audio analysis as an object state maker in bipolar disease using different classification algorithms such as RDA and SVM for classification of audio clips. we, divided the audio into frames for better classification and to extract the voice features of a clip.
Using the below-mentioned datasets from the table 2. We, classified datasets using different classification algorithms and achieved different accuracy for different types of audio. The accuracy from table 2 represents the increase in performance from area under curve value of 0.65±1.2 for maniac and 0.64±0.11 for depressed to the moderately higher area under curve value of 0.72±0.05 and 0.75±0.12, respectively. This excludes the results without noise and a two-way voice exchange of audio clips taken from user's smartphones. These algorithms can be compared with emotional analysis with Berlin and Spanish databases.

Table [2] : Accuracies of different classification algorithms in predicting bipolar disorder.

Dataset	Classification	Accuracy (%)
RECOLA	RDA Algorithm	65
SEWA	SVM Algorithm	63
BERLIN	RDA and SVM Algorithm	72
SPANISH	RDA and SVM Algorithm	70

V. FUTURE SCOPE

Identification of Bipolar disease in a particular person takes significantly more time compared to any other mental illness diseases and good care must be initiated by the management. Based on the past and present clinical results points a patient with bipolar disorder must be examined to clear doubts regarding early intervention and. It is hoped that early detection or identification of manic or depression will allow patients to take preventative measures. Efforts are being made to identify bipolar diseases at high risk.
The methods that are currently using by the examiners or doctors are not enough, as they must examine based on weekly or monthly interview calls. They currently examine users on their call and this is not enough to how a patient reacts to different factors. So, their personal calls are also recorded for research and improvement of control factors that give variations in subject symptomatology, communication pattern, and behavior. The application software that is developed for this study may adaptable and compatible with all devices.

VI. CONCLUSION

The methods that are currently using by the examiners or doctors are This literature survey paper presents a way to increase the performance of voice data collected from various mobile phones. It presents how different classification algorithms can classify different datasets with different accuracy and provides the best algorithm that provides more accuracy on a given one. This can be used for any health research organization for audio analysis of a mobile phone that wants to provide user-health based on noise-free and one-way data collected through a patient smart device. However, the data examined using the datasets and different algorithms are confined to recorded interview calls of people with bipolar disease and calls are taken every week. To know accurate symptoms their personal calls are also recorded to know factors of pitch variation concerning symptomatology, memory patterns and communication styles.

REFERENCES

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-iv disorders in the national comorbidity survey replication. Archives of general psychiatry. 2005; 62(no. 6):593–602. [PubMed: 15939837].
2. Angst J, Sellaro R, Angst F. Long-term outcome and mortality of treated versus untreated bipolar and depressed patients: a preliminary report. International Journal of Psychiatry in Clinical Practice. 1998; 2(no. 2):115–119. [PubMed: 24946291].
3. National Institute of Mental Health. [Accessed: September - 2015] Bipolar disorder in adults. http://www.nimh.nih.gov/health/publications/bipolar-disorder-in-adults/ Bipolar_Disorder_Adults_CL508_144295.pdf.
4. Hamilton, M. ECDEU Assessment Manual for Psychopharmacology, Revised Edition. National Institute of Mental Health; Rockville, MD: 1976. Hamilton depression scale; p. 179-92.
5. Young R, Biggs J, Ziegler V, Meyer D. A rating scale for mania: reliability, validity, and sensitivity. The British Journal of Psychiatry. 1978; 133(no. 5):429–435. [PubMed: 728692]
6. 6.Karam, Z.; Provost, E. Mower; Singh, S.; Montgomery, J.; Archer, C.; Harrington, G.; Mcinnis, M. Ecologically valid long-term mood monitoring of individuals with bipolar disorder using speech. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); May 2014; p. 4858-4862. 7.. LiKamWa, R.; Liu, Y.; Lane, ND.; Zhong, L. Moodscope: building a mood sensor from smartphone usage patterns. ACM International Conference on Mobile Systems, Applications, and Services; 2013; ACM; p. 389-402.
7. Osmani, V.; Maxhuni, A.; Grünerbl, A.; Lukowicz, P.; Haring, C.; Mayora, O. Monitoring activity of patients with bipolar disorder using smart phones. Proceedings of the International Conference on Advances in Mobile Computing & Multimedia; 2013; ACM; p. 85.
8. Sajana, T. & Narasingarao, M.R. 2017, "Ensemble framework for classification of imbalanced malaria disease data", Journal of Advanced Research in Dynamical and Control Systems, vol. 9, no. Special Issue 18, pp. 473-482.
9. Sajana, T. & Narasingarao, M.R. 2017, "Machine learning techniques for malaria disease diagnosis - A review", Journal of Advanced Research in Dynamical and Control Systems, vol. 9, no. Special Issue 6, pp. 349-369.
10. Razia, S. & Narasingarao, M.R. 2017, "A neuro computing frame work for thyroid disease diagnosis using machine learning techniques", Journal of Theoretical and Applied Information Technology, vol. 95, no. 9, pp. 1996-2005.
11. Anila, M. & Pradeepini, G. 2017, "Study of prediction algorithms for selecting appropriate classifier in machine learning", Journal of Advanced Research in Dynamical and Control Systems, vol. 9, no. Special Issue 18, pp. 257-268.
12. Bangare, S.L., Pradeepini, G. & Patil, S.T. 2017, "Brain tumor classification using mixed method approach", 2017 International Conference on Information Communication and Embedded Systems, ICICES 2017.
13. 14.Koteswara Rao Anne, Swarna Kuchibhotla and Hima Deepthi Vankayalapati, “Acoustic Modeling for Emotion Recognition," Springer Briefs In Electrical And Computer Engineering, Speech Technology, 2015.
14. Swarna Kuchibhotla, HD Vankayalapati, RS Vaddi, and KR Anne, “A comparative analysis of classifiers in emotion recognition through acoustic features," International Journal of Speech Technology,Springer Publishers,Volume 17 Issue 4, pages 401-408, December 2014.
15. Rao, K. R., & Josephine, B. M. (2018, October). Exploring the Impact of Optimal Clusters on Cluster Purity. In 2018 3rd International Conference on Communication and Electronics Systems (ICCES) (pp. 754-757). IEEE.

AUTHORS PROFILE

K. Naga Yaswanth is a student in department of Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, Vaddeswaram, AndhraPradesh

K. Samantha Roy is a student in department of Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, Vaddeswaram, AndhraPradesh

Ganesh Naga Sai Nerella is a student in department of Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, Vaddeswaram, AndhraPradesh

Dr Swarna Kuchibhotla is working as an Associate Professor in the department of Computer Science and Engineering at Koneru Lakshmaiah Education Foundation, Vaddeswaram, AndhraPradesh . Her main areas of research interests are Machine Learning , Speech processing emotional intelligence and Deep Learning.

Ms. B.Manjula Josephine working as an Assistant Professor,in the department of computer science and engineering Koneru Lakshmaiah Education Foundation,. Vaddeswaram. Andhra Pradesh. Had 2 years of teaching experience.Her main areas of research interest are Machine Learning and Computer Forensics