Depression Analysis of Voice Samples Using Machine Learning: A Review

Yashi Sharma, M. Tech Scholar, Department of Computer Science & Engineering, R.B.S. Engineering Technical Campus, Bichpuri, Agra, India.

Dr. Brajesh Kumar Singh, Professor, Department of Computer Science & Engineering, R.B.S. Engineering Technical Campus, Bichpuri, Agra, India.

Abstract

Depression as a phenomenon has evolved in multifaceted forms associated with individuals that have now led to complex hindrances in the overall wellbeing of a person. The pace with which depression has evolved in the sphere of life ranging from economic matters, family hassles or the havoc created by major diseases is a critical intruding factor in people's life than ever before.

Talking about the statistics, as per World Health Organization, India remains the sixth most depressive country in the world and the lack of objective assessment methods has led to a delay in proper treatment of people. This compared to other diseases where significant biomarkers/medical evaluations are available, there has been an increasing need to develop automated techniques for assessing psychiatric problems with several people.

This paper aims to analyze the scientific techniques and methods that have been under the lieu of development to predict the accuracy of depression in various people by assessing some notable features in the voice samples. This review paper incorporates data from the most scientific articles and organizations that have been a part in studying depression through Machine learning approaches.

Keywords

Depression, machine learning, machine learning algorithms and voice sample.

1. INTRODUCTION

Both anxiety and depression have been the emerging silent killers in people which might emerge as early as the age of 3-4 yrs.

Depression is a common illness which seem to be affecting more than 264 million people and these stats become alarming as the World Health Organization(WHO) states that nearly 800 000 people commit suicide every year due to depression.[1] The diagnosis of this rapidly spreading lifestyle disease becomes more difficult due to the pressure of social stigma, cost and availability of appropriate treatment. This is further aggravated with the unavailability of proper counselling and psychologist across the world.

Where nearly 20% of children suffer from this disorder in early childhood, this condition may impair the functioning and development of a child which may have serious implications later.[2].

The aforesaid risks coupled with societal burdens highlight the urgency of need in developing a method for early detection and diagnosis of depression as well as its originating symptoms.

Recent researches have tried figuring out characteristics from audio signal samples which allow the detection of signs leading to mental health diagnosis. Toanalyze these originating disorders from relevant data collected through audio samples a data driven technique such as machine learning can be the ideal way to make an appropriate conclusion from data.

Some researchers have authentically proved that people suffering from depression may speak in a different voice pattern that varies from the voice pattern of a healthy person. [3]. These voice samples studied under varied algorithms of machine learning not only possess a solution to the early detection of this problem but is also instrumental while preserving a person's privacy.

Figure 1 depicts the depression statistics in United States for varying age group.

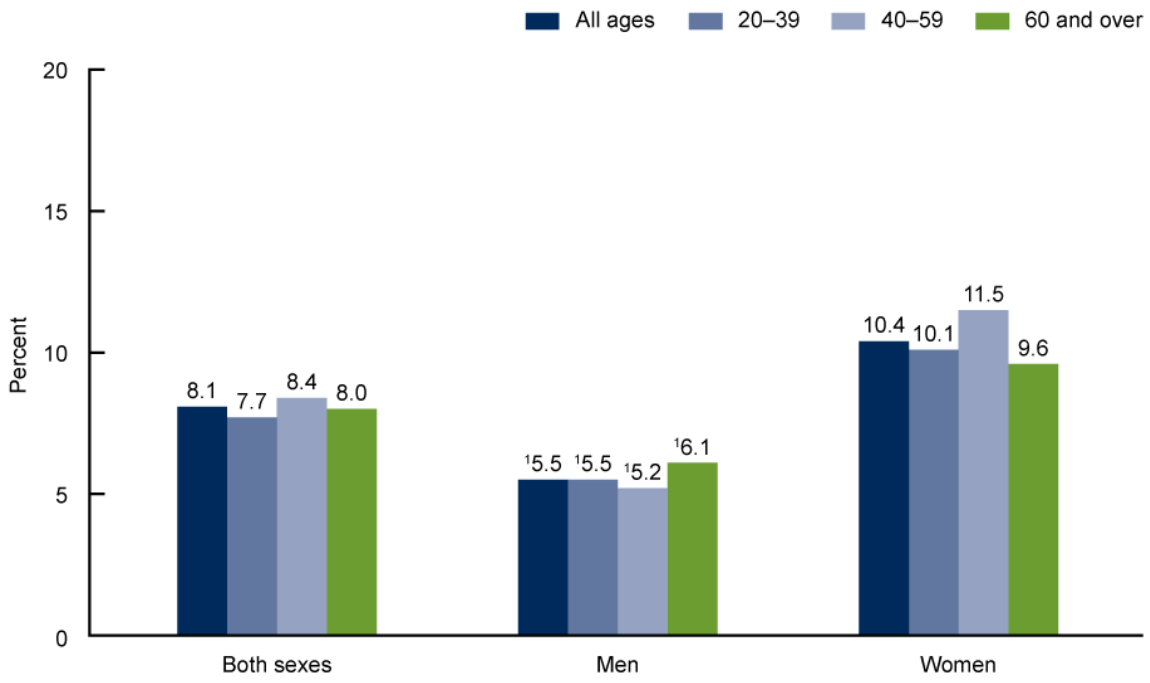

Figure 1: Depression Statistics in United States ,2013-2016 [4].

2. MACHINE LEARNING IN DEPRESSION ANALYSIS

Recent researches in the field of depression analysis have stated that depression can be figured out from speech with varying degrees [5]. Some researchers have been focused at the relations among level of depression and some prominent features that were noticed at a significant age. This conclusion was made from data samples collected by clinically depressive patients voice samples.

Features such as pitch of voice, loudness, frequency of speaking, pause taken in sentences and words, restlessness in voice, rate of speaking, vocal tract spectrum, etc.

All these listed features have been in studies as a part of human cognition in the past years.

Adding to this some new voice features like glottal waveform and Teager Energy Operation (TEO) which are presented in recent years and shown strong correlations to depression recognition [6]. Along with this, impact of acoustic features in depression has also been studied in some populations. While studying speaking style, it has been concluded that amongst and reading and spontaneous speech, the later one has shown more variability in people with depressive symptoms.
People who have been suffering from clinical depression have shown reduced vowel space while speaking sentences as compared to a normal person's way of speaking.

Voice signals have proven to be a remarkable information carrier in the assessing the mental health of the speaker. Various algorithms of machine learning approach have been the most generously used ones to evaluate depressive symptoms.

Researchers have been using various algorithms and modified them depending upon the problem statement. Of these algorithms, some approaches that have been widely used are:

Naïve Bayes algorithm: This algorithm can be used when the selected features in voice sample are independent from each other. Beinga model, which is based on prediction it has been in wide use to combine evidence collected from various features.

Decision Trees: They make segregation depending upon the feature values. Each node in a Decision tree symbols a feature with each division depicting the value which the node can have.

A decision tree partitions the data into subdivisions that may contain data with similar reoccurring values called split selection. This split selection finds an attribute and its corresponding splitting function.

Support Vector Machines: Support vector machines predict two different classes. It reduces the issue of over fitting by adjusting several features thus managing an excellent performance while feature selection from any sample.

The below figure 2 illustrates the accuracy of various machine learning algorithms while assessing depression.

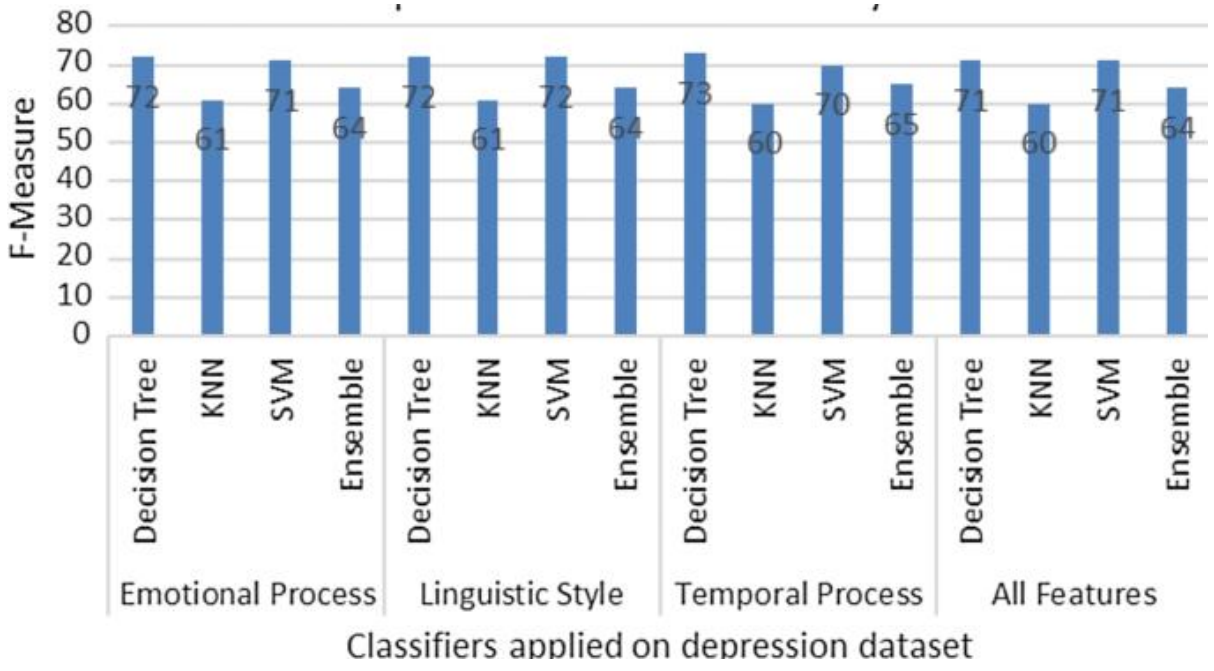

Figure 2: Depression Detection Accuracy

A study conducted on 33 volunteers from both genders had 11 healthy participants and 22 depressed participants. The depressive patients were diagnosed with major depressive disorder by a professional. All participants gave a written consent to be a part of the studies and data for depressive patients were collected from various hospitals.

Voice samples from this survey group was collected by a PCM recorder in WAV format. These patients were given a rating based on Self-reported Questionnaire (SRQ) for healthy patients and Hamilton Depression Rating scale was used for depressive patients.

The mean scores were calculated for patients in the age group for 28 to 40 yrs. Hamilton Depression scale predicted the mean score around 19.32.

Any unneeded voice signal was removed from the voice sample using audacity software. The vocal feature extraction from these recordings was done by GNU octave. Several features were extracted from these voice samples obtained which include features ranging from mean frequency range, spectral moments, amplitude variations to waveform length. [8]

Further to this machine learning algorithms like multilayer perceptron, random forest algorithm, decision tress, naïve Bayes and Support vector machines was used to classify the best algorithm for accessing results. The classification performance of various algorithms is shown in Table 1.

Table1: Classification Performance of various machine learning algorithms [8]

Model	Accuracy	Sensibility
MLP Model	81.2%- 81.9 %	81.6%-81.9%
Random Forest Model	84.2%-85 %	84.2%-84.9%
Decision Trees Model	74.2%- 74.6%	74.2%-74.7%

In addition to this, the study also observed a classification accuracy for both healthy as well as depressed individuals. The range of percentage of depression amongst the control group and healthy group can be clearly established as per Table 2.

Table 2: Confusion matrix for accuracy [8]

Group	Classified as Control	Classified as Depression
Control	88.25% 88.1-88.3 %	11.75% 11.5- 11.8
Depression	9.97% 9.6-9.97%	9.03%

3. COMPARATIVE ANALYSIS AND DISCUSSION

The analysis and methodology used in this study that various machine learning techniques can be successfully employed in concluding the accuracy for depression in voice samples [9,10].

However as different techniques are available and have been widely in use, out of these some techniques have given more accurate results then the others as shown in table 2.

Experiments that were conducted with fixed and adjustable settings showed varied results for varying algorithms. We can nearly conclude that this range of accuracy can predict outcomes for depressive as well as healthy patients.

However one limitation to these accuracy predictions may be factors that have an impact on the voice patterns of people which may include smoking as an history or any other medical limitation which might not be under control and can be a limitation in while predicting results [11,12].

The accuracy varied for different kernel algorithms and this variation existed between 67.69% to 89.14%. As per our review for various algorithms we can summarize the accuracy in the range as shown in bar graph in figure 3 and figure 4.

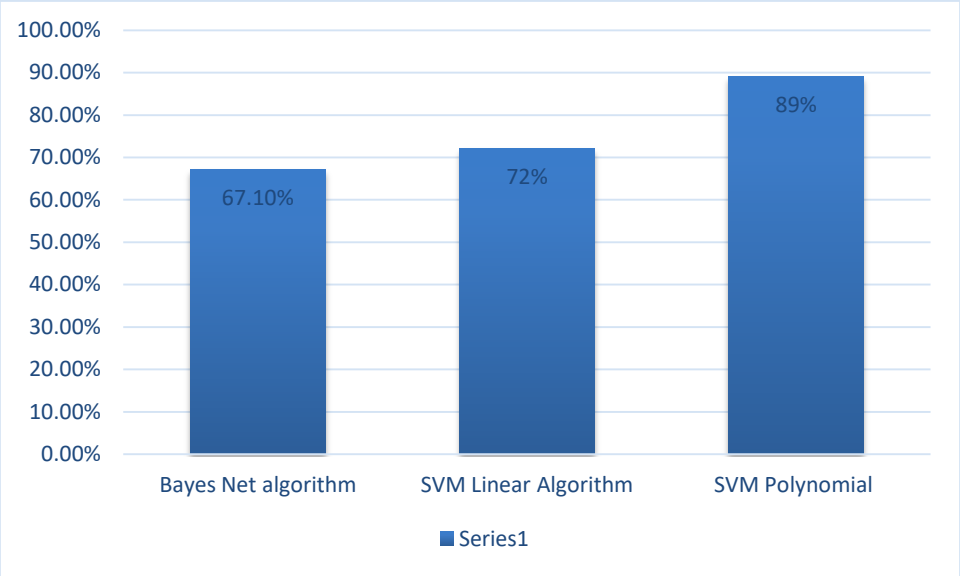

Figure 3: Bar chart depiction of accuracy of algorithms

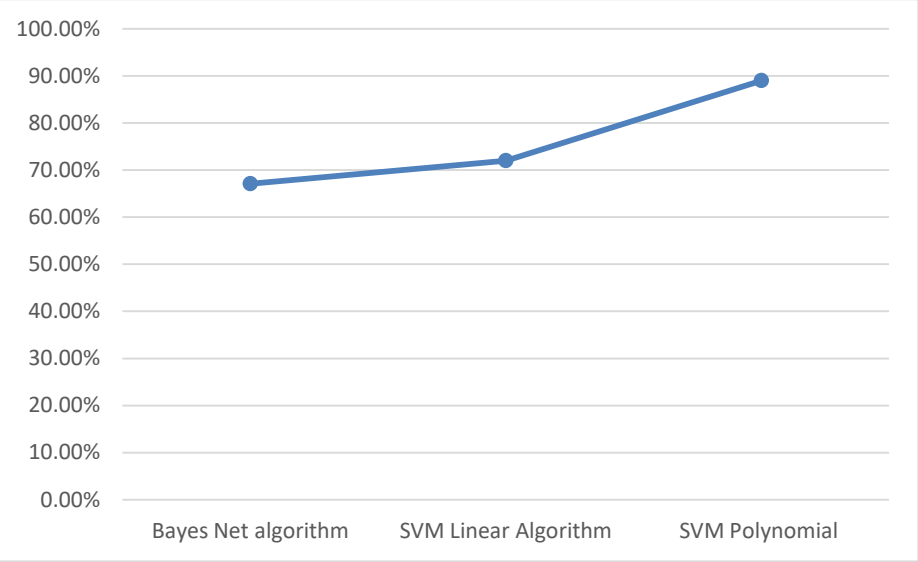

Figure 4: Line curve depiction of accuracy of algorithms

4. CONCLUSION

In this paper we briefed about how several machine learning techniques can analyze the anxiety patterns in different control groups.

The models obtained by applying machine learning techniques can efficiently and accurately outperform clinical evaluations that happen to predict the depressive symptoms in a patient.

The feasibility of these machine learning algorithms is also established as they are cost effective and non-invasive which means once a model has been designed for predicting depressive symptoms it can be used for a large group.

Further empirical observation on other data sets can provide improved outcomes on how different vocal properties in different data sets influence the performance and prediction accuracy of several automated classifiers in machine learning.

5. REFERENCES

1. https://www.who.int/news-room/fact-sheets/detail/depression.
2. J. L. Luby, A. C. Belden, J. Pautsch, X. Si, and E. Spitznagel,"The clinical significance of preschool depression: Impairment in functioning and clinical markers of the disorder," *J. Affect. Disorders*, vol. 112, no. 1–3, pp. 111–119, Jan. 2009.
3. Anastasia Pampouchidou, Panagiotis Simos, Kostas Marias, Fabrice Meriaudeau, Fan Yang, Matthew Pediaditis and Manolis, Tsiknakis, "Automatic Assessment of Depression Based on Visual Cues: A Systematic Review, vol. 10, no. 4,pp. 445-469, October-.December,2019.
4. https://www.cdc.gov/nchs/products/databriefs/db303.htm
5. K. E. B. Ooi, L.-S. A. Low, M. Lech, and N. Allen, "Early prediction of major depression in adolescents using glottal wave characteristics and teager energy parameters," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE* InternationalConference *on*, pp. 4613-4616,2012.
6. E. Moore, M. Clements, J. Peifer, and L. Weisser, "Investigating the role of glottal features in classifying clinical depression," in Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE, pp. 2849-2852,2003
7. Han, J., Pei, J., & Kamber, M. Data mining: concepts and techniques (Elsevier, 2011).
8. Espinola, C.W., Gomes, J.C., Pereira, J.M.S. *et al.* Detection of major depressive disorder using vocal acoustic analysis and machine learning—an exploratory study. *Res. Biomed. Eng.* **37,** 53–64, 2021.
9. Bzdok, D., & Meyer-lindenberg, A. (2018). Machine Learning for Precision Psychiatry: Opportunities and Challenges. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3, 223–230.
10. Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., S., & M., ... & Kerbrat, A. (2018). Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. Scientific Reports, 8(1), 1–17.
11. Darby, J. K., & Hollien, H. (1977). Vocal and Speech Patterns of Depressive Patiens. Folia Phoniatrica, 29, 279–291.
12. Hönig, F., Batliner, A., Nöth, E., Schnieder, S., & Krajewski, J. (2014). Automatic modelling of depressed speech: Relevant features and relevance of gender. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, (444), 1248–1252.