A Novel Framework for Speech-Based Detection of Schizophrenia using Machine Learning

Judy Flavia B, D. V. Vishnu Vardhan, G. Arvind Goud

Abstract*: Schizophrenia is a severe mental disorder that affects a person's thoughts, feelings and behavior. The disorder thus has serious impact on a person's personal and professional life. Traditionally the detection of schizophrenia is so far done with Electro Encephalography and MRI scans which make use of probabilistic methods. These methods are only useful when certain symptoms of the disorder are found. For early detection, a good method would be to use speech-based document using Conditional Random Fields algorithm. This method will use tagging of various speech components.

***Keywords*: Conditional Random Fields (CRF), Electro Encephalography (EEG), Magnetic Resonance Imaging (MRI), Schizophrenia.**

I. INTRODUCTION

Schizophrenia is a psychiatric disorder that is of great complex nature. Patients affected usually experience hallucinations, delusions and disorganized behavior. They also may exhibit problems with motor skills, in social life scenarios and their concentration on a subject may also be short-lived. It is a long standing problem that has received focus from various fields of science such as biomedical, neural etc. research into the topic has shown that early interference or diagnosis has led to people being affected recovering well. The procedure may involve certain medication and a generally peaceful environment is even more benefitting.

Though there are many such approaches, they are not very reliable. This is due to the fact that full working dynamics of the human brain are not totally observed.

The existing method of detecting schizophrenic condition is done using Electro Encephalography (EEG) and MRI scans.

The EEG technique is a non invasive procedure. Non-invasive procedures are those in which the skin is not cut or no contact made with the mucosa. It records the electric signals in the brain. The brain has billions of nerves which produce small electric signals when processing or carrying information amongst them. These signals form patterns called "brain waves". This is done by using small electrodes made of gold, silver etc. usually covered in silver chloride coating. Along with electrodes, wires are also attached to subject's head to detect brain waves and EEG machine amplifies signals and records them as wave patterns on monitor. Experts then derive conclusion from readings.

The MRI technique is based on magnetic field of small nuclei particles charged or in motion. The equipment then picks up the consequent small magnetic moments. MRI then measures how much water is present in different tissues inside brain. Since a large portion of our body is water, approximately at 65%, it is easier. Hydrogen is the atom precisely considered. When the human is under a magnetic field, the Hydrogen atoms in the brain align accordingly with the magnetic field. This behavior is called Larmor precession. The resonance signals then create images of the brain. Then they are compared with those of normal subjects using probabilistic algorithms to observe any disparities.

Their diagnostic results are inferior comparatively than those of human skilled professionals. Moreover, the mean and median of results vary and great importance should be given to design architecture and deep learning methods to get maximum possible results.

These methods are used when certain symptoms have been observed. So another method for detection of such conditions is done using oral document. Patients are required to produce a conversation about a series of events which is then recorded and worked upon. Natural Language Processing with CRF algorithm is used for effective outcomes.

Revised Manuscript Received on March 17, 2020.

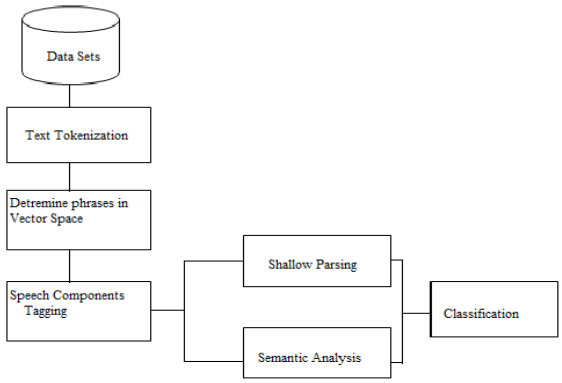
Judy Flavia B , Asst. Professor at SRM Institute of Science and Technology, Ramapuram, Department of Computer Science and Engineering, flavia4bruno@gmail.com.

D. V. Vishnu Vardhan, Student at SRM Institute of Science and Technology, Ramapuram, Department of Computer Science and Engineering, dvvvardhan@gmail.com.

G. Arvind Goud, Student at SRM Institute of Science and Technology, Ramapuram, Department of Computer Science and Engineering, arvindgoud142@gmail.com.

II. PROPOSED METHODOLOGY

A. Block Diagram



B. Building The Data Model

The first task includes getting the speech documents from the subjects which are then to be converted into the required text format using a dictation engine. This technology then records the speech in a sound waveform and applies signal processing to it. This helps to convert the sound into a compact representation. The Long Short term Memory –CRF (LSTM-CRF) network is used here. LSTM is a type of RNN. It can take inputs of larger values. The model helps to finish or predict the subject text. The data produced by subject are divided to fractions of milliseconds and the phonetics are then compared to existing letter phonetics. This way letters and then whole words are constructed and text can be prepared. The advantage is unlike regular RNN, it can learn long-term dependencies. This way, entire texts can be matched and data created.

C. Data Cleansing

The data acquired now is most likely composed of incomplete terms, useless phrases called noise and other inconsistencies. All this has to be removed for effective classification. For this purpose, text mining is applied to this data. It involves breaking up the data into sentences and then into words using grammar rules. It applies lexical rules to form meaningful sentences and phrase.

It includes removing unnecessary spaces, convert numerical stored as text into numbers, spell checking process and also proper formatting. This is achieved through use of tech like OpenRefine.

The processes are data reduction and data transformation among others. The aim here is to clear up the text of any discrepancies and then proceed onto tagging of the various terms or tokens in the text as speech components like verbs etc.

D. Shallow Parsing

Shallow parsing involves 2 steps, first identifying the various tokens as respective speech components and second, grouping them together to make sense. The difference between speech components tagging tokens and shallow parsing is that the former only detects individual parts of sentence but the latter detects and joins certain parts to form sub parts of a sentence. This leads to giving a higher sense of meaning to the sentence when coupled with machine learning.

E. Semantic Analysis

Semantic analysis refers to the process of checking if the text or part of it is grammatically correct or not. It takes into consideration the various rules like active to passive and vice-versa, tenses of speech etc. It also considers idioms and clauses and works on them to get their meaning. It will consider the many rules like synonyms, metaphors etc.

F. classification

The resulting texts are then compared with those of normal people to detect and classify subjects as schizophrenic and non-schizophrenic.

G. algorithm

The proposed system is based on the Speech Components Tagging from document characterization. This system automatically extracts and counts the lexical features and document vectors of speech components tags verbal content generated by schizophrenic patients and non schizophrenic patients.

Conditional Random Fields is a natural Language Processing algorithm in Machine Learning that can be used for tagging the speech components. It distinguishes the various components as verbs, nouns, adverbs etc. The challenge here is the same word can be used as a verb or as a noun in different contexts by the patients. For Example, in the sentence “Answer the question”, Answer is a verb or an action. In “Give me the answer”, Answer is a noun used as an alternative word for reply.

The system has two feature representations:

1. SC features
2. Meta-SC features

Speech Components (SC) will categorize words based on their usage in a phrase or in a sentence. The SC categorizes words as nouns, pronouns, adverbs, verbs, adjectives, prepositions, interjections and conjunctions.

WRC-CDRC is a method to allocate weight for words or tokens in a data document. It comprises of:

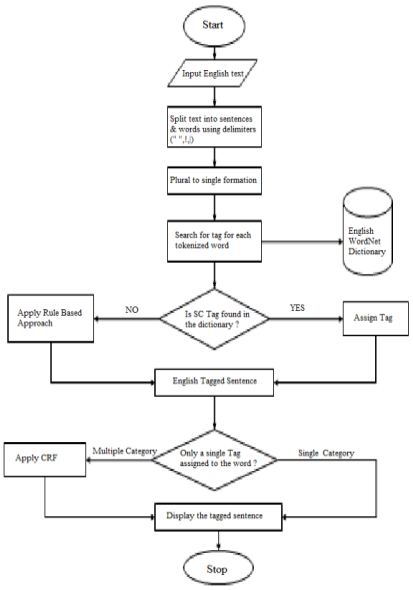
1. Word Recurrence Count (WRC) and
2. Contrary Document Recurrence Count (CDRC).

Their respective weights allow for evaluating how greatly they affect the sentence and its meaning. They are computed as:

WRC (w) = (Times term ‘w’ recurs in document) / (sum of terms in document).

CDRC (w) = (Sum of all documents / no. of documents with ‘w’ in them).

III. FLOWCHART



IV. RESULT ANALYSIS

Research on link between language and schizophrenia using Machine Learning is rare to come by. The model will point out any disorganized way of talk or other related symptoms during many stages of the disorder. Differences in grammar context were also found. The linguistic characteristics thus allow for good detection. Proposed system has advantages like hidden patterns in the document can be clearly identified. Also the linguistic analysis is also good.

V. CONCLUSION

The precise intention of our proposed model is to distinguish schizophrenic from non schizophrenic by making use of verbal document generated from the particular person. We utilized NLP to automatically evaluate the recordings of schizophrenic patients and healthy subjects. The schizophrenia patients display substantial linguistic divergences associated with non schizophrenia patients.

Proposed depictions for data transformed from oral speech are put up for comparison against standard characterization like TF-IDF. It initially proposed for record recuperation and now is adequately used for requesting content across different spaces other than its general idea. As a general finalization, we predict that system will help in successfully differentiating the subjects. Another big factor is the data were gotten from etymological endeavor, thus making data homogeneous.

The proposed system does encourage and present an essential phase towards our complete aim of establishing automated systems to aid clinical diagnosis and understanding of schizophrenia. Our system can be used by people with basic skills of operation. It is also cost ineffective.

FUTURE WORK

The system would work more efficiently with added linguistic features to work upon. The system fails to give accuracy in case of young children whose linguistic skills would mostly not be developed as yet. Also people who stammer or people with low or no proficiency in the language of the text are prone to be wrongly classified as being schizophrenic. Further leniency in the system in such cases must be provided. Also support for multiple languages can be incorporated. Collecting more data will be a step in the right direction for improved classification among the texts obtained.

REFERENCES

1. Christ Devia, Member, IEEE, Rocío Mayol-Troncoso, Javiera Parrini, Gricel Orellana, Aída Ruiz, Pedro E. Maldonado& José Ignacio Egaña, “EEG Classification During Scene Free-Viewing for Schizophrenia Detection “,IEEE Vol 27 June,2019.
2. Bo Yang , Yuan Chen, Quan-Ming Shao, Rui Yu, Wen-Bin Li, Guan-Qi Guo, Jun-Qiang Jiang, and Li Pan,” Schizophrenia Classification Using fMRI Data Based on a Multiple Feature Image Capsule Network Ensemble”, IEEE ACCESS Vol.7, August,2019.
3. Fatemeh Alimardani, Jae-Hyun Cho, Reza Boostani, Han-Jeong Hwang ,” Classification of Bipolar Disorder and Schizophrenia Using Steady-State Visual Evoked Potential Based Features”, IEEE Vol. 6,2018.
4. Min Wang , Student Member, IEEE, Heba El-Fiqi,Member, IEEE, Jiankun Hu , Senior Member, IEEE, and Hussein A. Abbass , Senior Member, IEEE,” Convolutional Neural Networks Using Dynamic Functional Connectivity for EEG-Based Person Identification in Diverse Human States”, IEEE VOL. 14, NO. 12, DECEMBER 2019.
5. Luciano Comin Nunes, Placido Rogerio Pinheiro (Member, IEEE), Mirian Caliope Dantas Pinheiro, Marum Simao Filho, Rafael Espindola Comin Nunes and Pedro Gabriel Caliope Dantas Pinheiro,” Automatic Detection and Diagnosis of Neurologic Diseases ”, IEEE Vol.27,2019.
6. Jiahui Qiu ,Member, IEEE , Yangming Zhou ,Member, IEEE ,Qi Wang,Member, IEEE, Tong Ruan, and Ju Gao ,“ Chinese Clinical Named Entity Recognition Using Residual Dilated Convolutional Neural Network With Conditional Random Field “,IEEE VOL. 18, NO. 3, JULY 2019.
7. Manohar Latha , Ganesan Kavitha ,” Detection of Schizophrenia in brain MR images based on segmented ventricle region and deep belief networks “,Neural Computing and Applications, Springer, Feb 2018 .
8. Indranath Chatterjee, Manoj Agarwal, Bharti Rana, Navin Lakhyani, Naveen Kumar,” Bi-objective approach for computer-aided diagnosis of schizophrenia patients using fMRI data ”, Multimed Tools Appl, Springer, AprilL, 2018.
9. Khalil Khan, Nasir Ahmad, Farooq Khan, Ikram Syed,” A framework for head pose estimation and face segmentation through conditional random fields”, Springer, May 2019.
10. Bernis Sutcubasi, Sinem Zeynep Metin, Turker Tekin Erguzel, Baris Metin, Cumhur Tas, Mehmet Kemal Arikan, Nevzat Tarhan, ”Anatomical connectivity changes in bipolar disorder and schizophrenia investigated using whole-brain tract-based spatial statistics and machine learning approaches”, Neural Computing and Applications, Springer Jan 2019.

AUTHORS PROFILE

Judy Flavia B , Asst. Professor at SRM Institute of Science and Technology, Ramapuram, Department of Computer Science and Engineering, flavia4bruno@gmail.com.

D.V.Vishnu Vardhan, Student at SRM Institute of Science and Technology, Ramapuram, Department of Computer Science and Engineering, dvvvardhan@gmail.com.

G.Arvind Goud, Student at SRM Institute of Science and Technology, Ramapuram, Department of Computer Science and Engineering, arvindgoud142@gmail.com.