Speech-Based Depression Diagnosis

Team IRIS

A. Team Members:

1. Rohan Pillai
2. Niraj Anil Babar
3. S. Loganyaa

B. Abstract:

Depression is a common mental disorder affecting 1 in 7 people worldwide. With limited access to qualified mental health professionals in India and around the world, it has become increasingly hard for people to understand their emotions and seek help when required.

The aim of our project is to create a system to aid doctors and patients in the screening of depression using machine learning and digital signal processing. Our system consists of 3 tests:

Test 1:

The first is a questionnaire to establish the patient's general behavior and mood. Researchers have discovered symptoms of depression that precede physical manifestation, which include insomnia, loss of weight, lack of concentration, and fatigue. Based on these symptoms we ask the patient a series of questions, these questions serve as a screening criterion, and if the patient shows any of the above-mentioned symptoms we proceed to the next test.

Each question has 4 options ranging from None of the symptoms observed to the symptom observed nearly every day. This test is based on the PHQ-9 , a popular depression screening test used by psychiatrists.

Test 2:

Our second and main test identifies the signs of depression from a person's speech. Studies have identified that depression can change the way we talk: our speech becomes flatter and more monotone and our frequency changes. Our system does not rely on the words spoken but rather on the acoustic features of the voice sample. Here we have used a CNN model to classify if the person shows signs of depression. In this, we convert the audio samples into its spectrograms. A spectrogram is a visual way of representing the signal strength over time at various frequencies. There are 2 dimensions to the graph with a third dimension represented by colors. It is an intensity plot of the Short-Time Fourier Transform.

We are training the model on the spectrograms of audio samples. While testing we ask the user to record his/her audio and then convert it to its spectrogram and test it.

Test 3:

Our final test is an (optional) social media screen. Depression, especially in teens and young adults is a direct result of cyberbullying. Further millennials tend to be more expressive about their emotion on social media compared to previous generations.

We leverage this by scraping the last 10 tweets from the patient and scanning them for signs of emotional distress. This includes tweets talking about anxiety, sleep disorders, or sadness.

Based on the above tests we generate a score out of 10. Based on the score we can advise the user on the next step i.e whether he's healthy or he needs to visit a qualified mental health professional. Further, we can plot a graph of these scores

against time. This graph can be used by the doctor to examine whether the patient is improving/deteriorating and how he is responding to treatment and medication.

C. Tech Stack:

- Librosa
- Flask
- Text-blog
- Tensorflow
- Keras
- CNN
- Sk-learn
- Numpy
- Pandas
- HTML /CSS

D. References:

1. Stephen, Jini & P., Prabu. (2019). Detecting the magnitude of depression in Twitter users using sentiment analysis. International Journal of Electrical and Computer Engineering (IJECE). 9. 3247. 10.11591/ijece.v9i4.pp3247-3255.

2. LangHe CuiCao . Automated depression analysis using convolutional neural networks from speech. https://doi.org/10.1016/j.jbi.2018.05.007

-FIN-