Title: Depression Analysis and Detection System Based on Audio Data

Team Information

Charlie(Shan Jiang), sjiang24, 50262137

Note: even though I’m myself, I will be mostly using “We”

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Introduction

Background

Depression is a widespread mental health issue affecting millions of people worldwide. According to the World Health Organization, depression is a leading cause of disability globally, impacting individuals' daily lives and overall well-being. Early detection and continuous monitoring are crucial for effective management and treatment of depression. Traditional methods of diagnosing depression often rely on subjective self-reports or clinical interviews, which can be time-consuming and prone to bias. These methods require significant time and effort from both patients and healthcare professionals, which can lead to delays in diagnosis and treatment. Symptoms of depression symptoms can vary from mild to severe and can appear differently in each person. These symptoms can include:

* Feeling sad, irritable, empty and/or hopeless.
* Losing interest or pleasure in activities you once enjoyed.
* A significant change in appetite (eating much less or more than usual) and/or weight (notable loss or gain unrelated to dieting).
* Sleeping too little or too much.
* Decreased energy or increased tiredness or fatigue
* Increase in purposeless physical activity (e.g., inability to sit still, pacing, handwringing) or slowed movements or speech that are severe enough to be observable by others.
* Feeling worthless or excessively guilty.
* Difficulty thinking or concentrating, forgetfulness, and/or difficulty making minor decisions.
* Thoughts of death, suicidal ideation, or suicide attempts.

Depression can happen to anyone. People who have lived through abuse, severe losses or other stressful events are more likely to develop depression. Women are more likely to have depression than men.

An estimated 3.8% of the population experience depression, including 5% of adults (4% among men and 6% among women), and 5.7% of adults older than 60 years. Approximately 280 million people in the world have depression (1). Depression is about 50% more common among women than among men. Worldwide, more than 10% of pregnant women and women who have just given birth experience depression (2). More than 700 000 people die due to suicide every year. Suicide is the fourth leading cause of death in 15–29-year-olds. From these, we can see the impacts depression has on varies settings, therefore, identifying depression is very important and we should take extra care of our health.

Motivation

With the advancement of machine learning and natural language processing technologies, it is possible to develop systems that automatically detect depression using audio data. Audio data provides a rich source of information that can be analyzed to detect changes in speech patterns, tone, and other acoustic features associated with depression. These systems can provide timely and objective assessments, helping individuals and healthcare professionals manage depression more effectively. Automated systems can reduce the burden on healthcare systems, provide continuous monitoring, and offer early intervention opportunities.

Project Overview

This project aims to develop a depression analysis and detection system based on audio data. The system utilizes machine learning models to analyze audio recordings, predict depression levels, and provide relevant feedback. The system is developed using PyQt5 for the user interface, torch for building machine learning models, and mysql.connector for database interactions. By integrating these technologies, we aim to create a comprehensive tool that is user-friendly, efficient, and accurate in detecting depression. When considering the language we want to use for voice inputs, since I’m a Chinese international student and both our TA and Instructor have Chinese Backgrounds, using Chinese as voice inputs is not a problem in understanding. However, if there are any concerns regarding this, I’m happy to make modifications, but right now, let us just stay in Chinese for voice inputs. I have downloaded some voice inputs from the internet for testing and analysis purposes, we can also record our voice from the app so that it can be analyzed as well, the result is simply returning 1 if the system detected that the user has depression and 0 if not, binary classification made us both simple and effective.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Related Work

Related Work Summary

Several studies have explored the potential of using audio features and machine learning for mental health assessment. For example, one study indicates that specific audio features can reflect the severity of depression, such as changes in pitch, tone, and speech rate. Another study uses deep learning techniques to classify audio recordings to detect depression and has achieved promising results. These studies provide a foundation for this project, demonstrating the potential of audio-based depression assessment systems. Additionally, previous research has highlighted the importance of feature selection and model architecture in improving the accuracy and reliability of these systems. We have included some related work links in the reference section.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Project Description

System Overview

The system consists of four main components: user interface, audio processing module, depression detection model, and database management. Users can record or upload audio files through the interface. The system will analyze the audio, output depression level assessments, and provide corresponding suggestions. Each component plays a critical role in ensuring the system's functionality and user experience.

User Interface

The user interface is developed using PyQt5 and includes the following main features:

Audio Recording and Uploading: Users can either record new audio samples directly through the interface or upload pre-recorded audio files. This flexibility allows users to provide data in the most convenient manner for them.

Questionnaire Function: In addition to audio analysis, users can complete a standardized depression questionnaire. This dual approach combines subjective self-reporting with objective audio analysis. The questionnaire consists of 20 questions, it serves as a survey in combination with audio analysis, by asking the users rate their respective levels on how sad they are currently, how sad they are in the past, how much food do they eat, do they feel that their mind is congested or not, etc., we can have a broader understanding of their issues, thus actions can be taken based on them.

Display of Depression Level Assessment Results: The interface provides a clear and concise display of the assessment results, including visual indicators of depression levels(for example, mild depression, moderate depression, severe depression, etc).

Audio Processing Module

This module is responsible for converting user-recorded or uploaded audio files into a format suitable for the model. The main steps include:

Converting MP3 Files to WAV Format: Ensures compatibility with the audio processing pipeline and standardizes the input format.

Using the python speech recognition Library for Speech-to-Text Conversion: Converts spoken words into text, which can be analyzed for linguistic features.

Extracting Audio Features for Model Use: Key features such as pitch, tone, energy, and speech rate are extracted to be used as input for the machine learning model. Feature extraction is crucial for capturing the nuances of speech that may indicate depression.

Depression Detection Model

We use a pre-trained AudioClassifier model to classify audio data. The model uses 13 input features, processes them through two hidden layers, and outputs the prediction results. The model is trained and evaluated using torch. The architecture of the model is designed to capture complex patterns in the audio data that are indicative of depression. The training process involves using a large dataset of labeled audio samples to fine-tune the model's parameters for accurate prediction.

Database Management

The system uses mysql.connector to interact with a MySQL database, storing user audio data and assessment results. The main functions include:

Connecting to the Database: Establishes a secure and efficient connection to the MySQL server.

Storing and Querying Audio Features and Assessment Results: Ensures that all data is stored in an organized manner and can be easily retrieved for analysis and reporting.

Matching Similar Questions: Enhances the user experience by providing relevant feedback based on previous assessments and similar cases.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Experiments

Experimental Setup

We conducted several experiments to validate the effectiveness of the system. The main experiments included:

Evaluating the Accuracy of Depression Detection Across Different Audio Samples: Testing the system with a diverse set of audio recordings to assess its generalizability and accuracy.

Comparing Questionnaire Results with Audio Analysis Results: Cross-validating the system's predictions with traditional self-report measures to ensure consistency and reliability.

Experimental Results

The experimental results show that the system has a high accuracy in predicting depression levels. Additionally, there is a significant correlation between the questionnaire results and audio analysis results, confirming the system's effectiveness. The system was able to detect subtle changes in speech patterns that corresponded with self-reported depression levels, demonstrating its potential for real-world application.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Conclusion

This project developed a depression analysis and detection system based on audio data. By utilizing machine learning models and database management technologies, the system can provide timely and objective depression level assessments. The experimental results validate the system's effectiveness, demonstrating its potential for practical application. Future work may involve refining the model, expanding the dataset, and integrating additional features to enhance the system's accuracy and usability.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

References

Abdul Rahman, H.; Kwicklis, M.; Ottom, M.; Amornsriwatanakul, A.; H. Abdul-Mumin, K.; Rosenberg, M.; Dinov, I.D. Machine Learning-Based Prediction of Mental Well-Being Using Health Behavior Data from University Students. Bioengineering 2023, 10, 575. https://doi.org/10.3390/bioengineering10050575

Nasser Ghadiri, Rasoul Samani, Fahime Shahrokh, Integration of Text and Graph-based Features for Detecting Mental Health Disorders from Voice. https://arxiv.org/abs/2205.07006

Khoo, L.S.; Lim, M.K.; Chong, C.Y.; McNaney, R. Machine Learning for Multimodal Mental Health Detection: A Systematic Review of Passive Sensing Approaches. Sensors 2024, 24, 348. https://doi.org/10.3390/s24020348

Iyortsuun, N.K.; Kim, S.-H.; Jhon, M.; Yang, H.-J.; Pant, S. A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis. Healthcare 2023, 11, 285. <https://doi.org/10.3390/healthcare11030285>

Python | Introduction to PyQt5, https://www.geeksforgeeks.org/python-introduction-to-pyqt5/

Torch, https://pytorch.org/docs/stable/torch.html

Add userbase library dir to windows dll search path, <https://github.com/pytorch/pytorch/commit/fdfef759a676ee7a853872e347537bc1e4b51390>

Depressive disorder (depression). <https://www.who.int/news-room/fact-sheets/detail/depression>

Chapter 1 Introduction to MySQL Connector/Python. <https://dev.mysql.com/doc/connector-python/en/connector-python-introduction.html>.

MySQL Connectors, <https://www.mysql.com/products/connector/>

David Amos, The Ultimate Guide To Speech Recognition With Python, <https://realpython.com/python-speech-recognition/>

How to Install FFmpeg on Windows 11. https://www.youtube.com/watch?v=22vmzTs5BoE