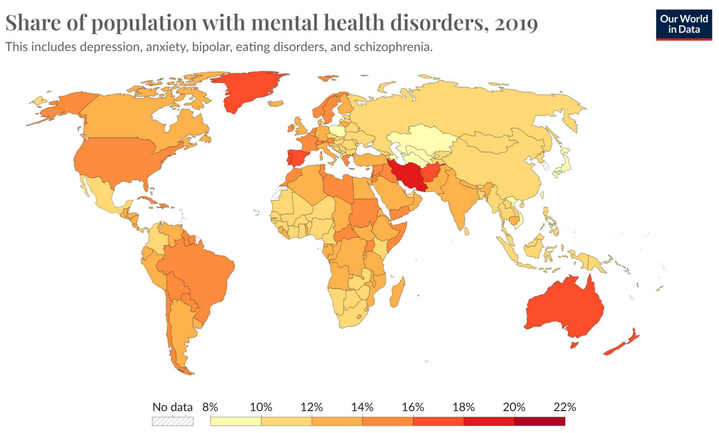
# A Machine-Learning Model for Detecting Depression, Anxiety, and Stress from Speech

**Mashrura Tasnim,** Ramon Diaz Ramos, Eleni Stroulia and Luis A. Trejo

## Motivation

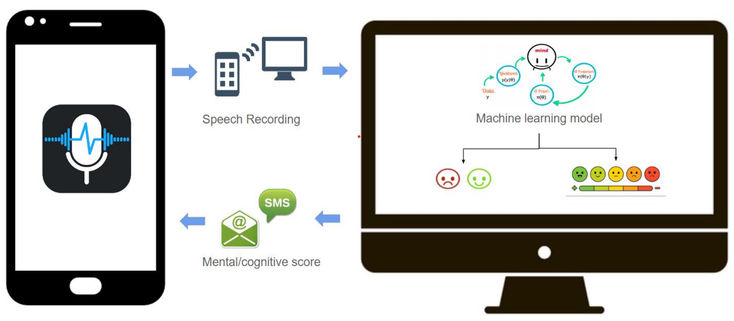
"A mental disorder is characterized by a clinically significant disturbance in an individual's cognition, emotional regulation, or behavior. " -WHO




- 1 in every 8 people live with a mental disorder
- Most patients cannot access treatment
- Symptoms can be detected from voice

## Objective

To develop a practical system that can capture the audio of the users' voices and analyze it to estimate their depression, anxiety and stress level




## Contribution

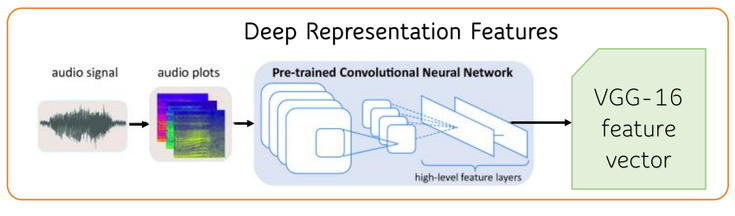
1) Curated YouthDASS dataset

- Multilingual longitudinal speech corpus to study depression, anxiety, and stress
- Labelled with DASS-21 scores

2) Validated participants' adherence to study protocol

3) Explored the effectiveness of our speech processing pipeline on the YouthDASS data set

- VGG-16 features were extracted from speech spectrograms



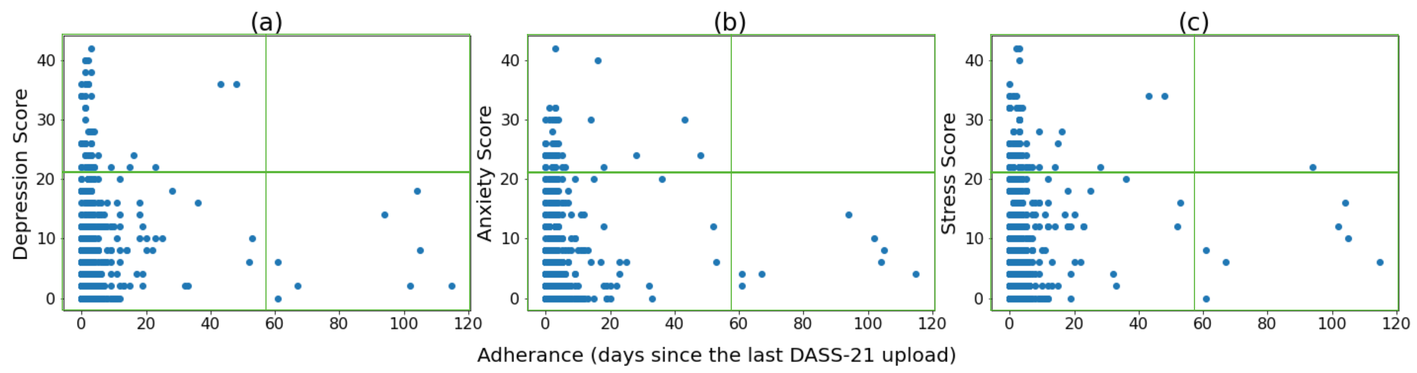

- 2 layer 1-D CNN model was trained for predicting depression, anxiety and stress severity

## YouthDASS Dataset

| Longitudinal depression, anxiety and stress dataset | |
|---|---|
| **Language** | English and Spanish |
| **# Participants** | 40 (26 Spanish, 14 English) |
| **# Recordings** | 1,049 (838 Spanish, 211 English) |
| **Avg #samples per participant** | **26** |
| **Average DASS-21 score** | 15.3 (SD = 12.3) |

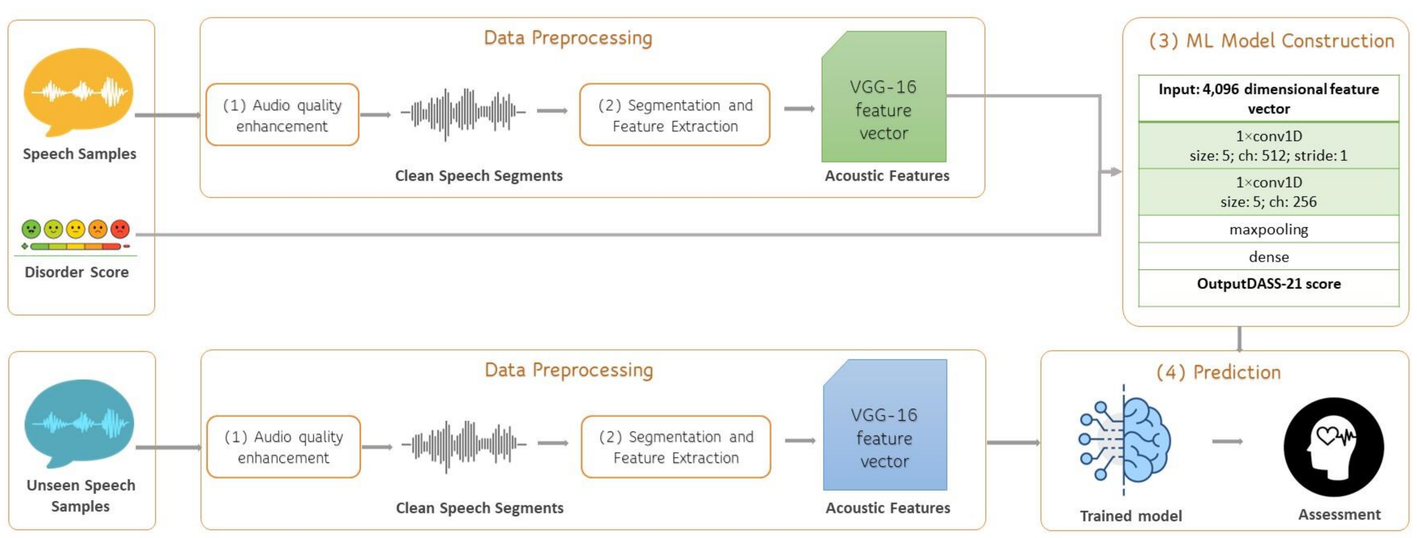
| DASS-21 scores in YouthDASS | Depression | Anxiety | Stress |
|---|---|---|---|
| **Threshold** | 9 | 7 | 15 |
| **Mean** | 5.94 | 4.79 | 7.12 |
| **Std. Dev** | 7.55 | 6.90 | 8.20 |
| **Median** | 4 | 2 | 4 |

## Participants' Adherence To Study Protocol




Participants with lower DASS-21 scores adhered better to the data collection procedure, designed as a routine activity to be performed twice a week

## Method




## Results

| | Citation | Method | RMSE | NRMSE |
|---|---|---|---|---|
| Depression | Kim et al., 2013 | Deep CNN on log-Mel spectrograms | - | 0.205 |
| | **Ours** | **1-D CNN trained on VGG-16** | **7.09** | **0.169** |
| Anxiety | Fatima et al. 2021 | MLP on sentimental BOW extracted from 142 suicide notes | - | 0.089 |
| | **Ours** | **1-D CNN trained on VGG-16** | **7.69** | **0.183** |
| Stress | Fatima et al., 2021 | MLP on sentimental BOW extracted from 142 suicide notes | - | 0.103 |
| | **Ours** | **1-D CNN trained on VGG-16** | **8.40** | **0.200** |

Our proposed 1-D CNN model using **acoustic features exclusively** achieved better NRMSE than state-of-the-art models in **depression prediction**; and performed competitively with acoustic and linguistic models in predicting anxiety and stress

## Key Takeaways

1. **YouthDASS dataset** captures valuable information on the post-pandemic effect on the mental health of youths
2. Our proposed **CNN model** demonstrates competitive performance to the state-of-the-art acoustic and linguistic models in predicting depression, anxiety, and stress severity

UNIVERSITY OF ALBERTA
Tecnológico de Monterrey