# Depression Detection

- Week 8

Name: Sinchana Kumbale

University: Manipal Institute of Technology

Internship Period: Jan – July 2024

Undergraduate 4th year

Duration of the presentation: ~10 minutes

# Agenda

1. **Balancing Classes for Anxiety Detection Code**

2. **Subset Creation Based on Emotion Relevance**

   1. Methodology and Results

   2. Comparison with the prior topic-based consolidation

3. **Insights from paper - Acoustic and Landmark Event-Based Features**

4. **Literature Organization – Mind Map**

5. **Tentative plan for next week**

# Balancing Classes for Anxiety Detection Code

Table 1: The results of running methodology [1] locally on DAIC-WOZ ( macro avg metric / micro avg metric/ upsampled)

| Method | Precision | Recall | F1 | AUC-ROC |
|---|---|---|---|---|
| Raw Text | 0.59 / 1.0 / 0.67 | 0.51 / 0.14 / 0.48 | 0.60 / 0.25 / 0.56 | 0.65 |
| Pre-processed Text | 0.66 / 0.50 / 0.60 | 0.57 / 0.21 / 0.54 | 0.56 / 0.30 / 0.57 | 0.64 |
| Emotional Text | 0.68 / 1.0 / 0.62 | 0.61 / 0.21 / 0.58 | 0.60 / 0.35 / 0.60 | 0.68 |

[1] Detecting anxiety from short clips of free form speech (Agarwal et al, 2023)
| Paper | Code
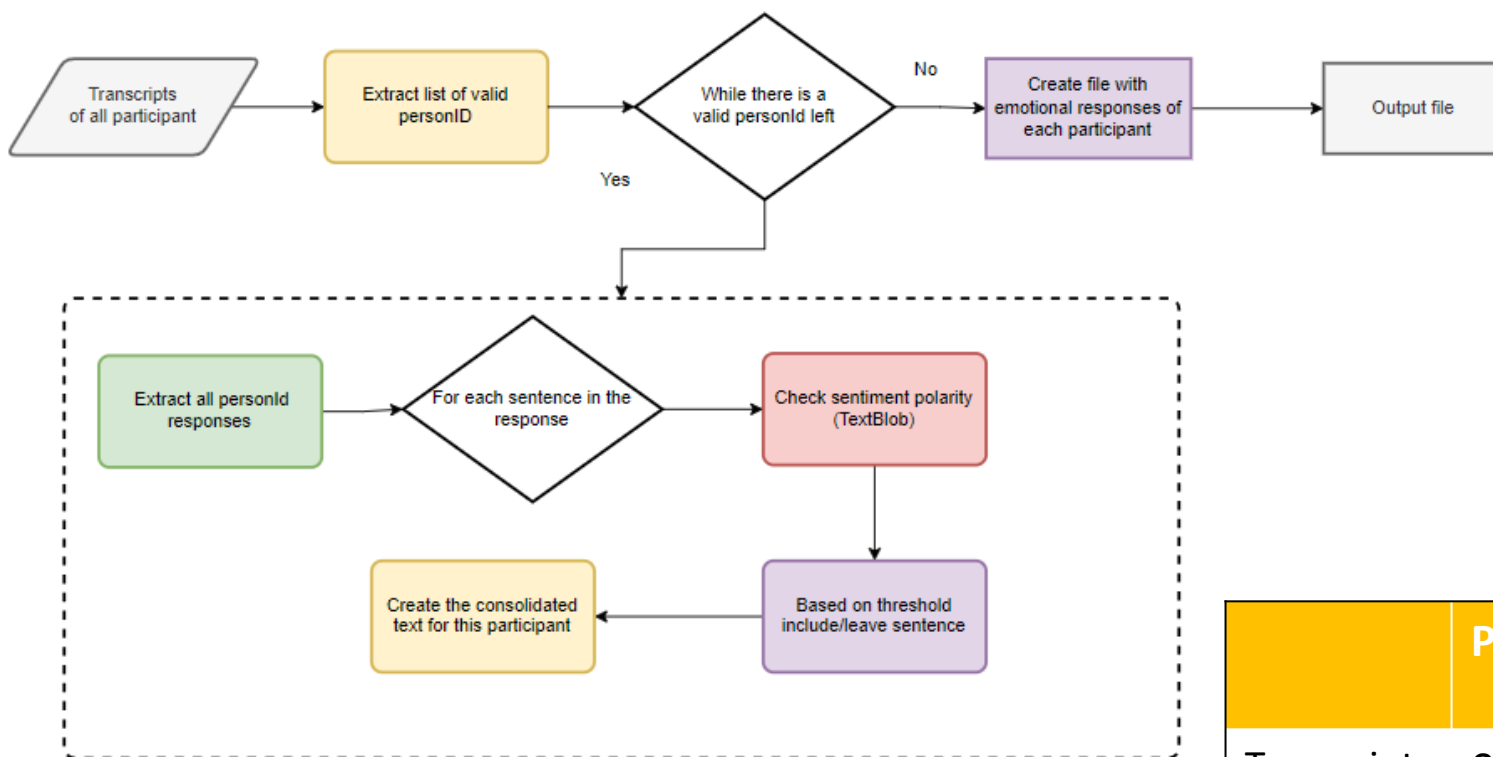
# Subset Creation Based on Emotion Relevance



Fig 1: Methodology to create a subset containing only emotional responses of each participant

Table 2: Results of running Yuxin's code locally

| | Precision | Recall | F1 | Accuracy |
|---|---|---|---|---|
| Transcripts | 0.63 | 0.67 | 0.65 | 0.66 |
| Question based | 0.62 | 0.64 | 0.63 | 0.62 |
| Emotion Based | 0.83 | 0.58 | 0.68 | 0.73 |

# Subset Creation Based on Emotion Relevance

Table 3: Comparing Question based consolidation and emotional consolidation

| Method | Pros | Cons |
|---|---|---|
| Question (Topic) based consolidation | 1. Condensing a large transcript to only relevant and informative details | 1. Requires some level of manual selection<br>2. Potentially missing out important information of a participant |
| Emotional responses only | 1. Condenses transcripts by removing all neutral sentences<br>2. Can be easily applied to other datasets with no changes | 1. Might miss out on context that could influence the emotional responses |

# Insights from paper



Fig 2: Method followed by [2]
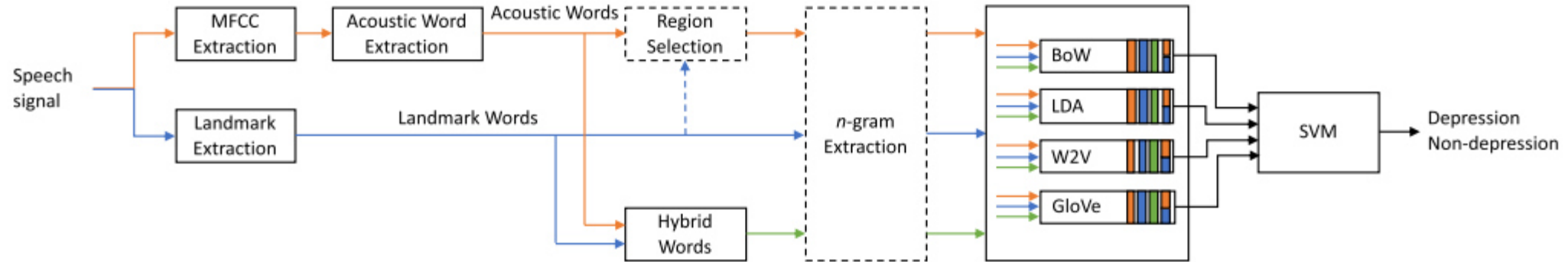
## TABLE I
### DESCRIPTION OF THE SIX LANDMARKS INVESTIGATED

| Landmark | Description |
|---|---|
| g | sustained vibration of vocal folds starts (+) or ends (−). |
| p | sustained periodicity begins (+) or ends (−) |
| s | releases (+) or closures (−) of a nasal |
| f | frication onset (+) or offset (−) |
| v | voiced frication onset (+) or offset (−) |
| b | onset (+) or offset (−) of existence of turbulent noise during obstruent regions |

[2] Natural Language Processing Methods for Acoustic and Landmark Event-Based Features in Speech-Based Depression Detection (Huang et al, 2020) | Paper

6

# Insights from paper

| | F1(D) | F1(H) | Acc. | Conf. Mat. |
|---|---|---|---|---|
| AVEC 2016 (A) [47] | 0.41 | 0.58 | 51.4% | $\begin{bmatrix} 12 & 16 \\ 1 & 6 \end{bmatrix}$ |
| AVEC 2016 (A+V) [47] | 0.58 | 0.86 | 77.1% | $\begin{bmatrix} 22 & 6 \\ 2 & 5 \end{bmatrix}$ |
| DepAudioNet (A) [54] | 0.52 | 0.70 | 65.7% | $\begin{bmatrix} 15 & 13 \\ 0 & 7 \end{bmatrix}$ |
| Audio + Gender [55] | 0.59 | 0.87 | - | - |
| AW | 0.636 | 0.833 | 77.1% | $\begin{bmatrix} 20 & 8 \\ 0 & 7 \end{bmatrix}$ |
| LW | 0.462 | 0.682 | 60.0% | $\begin{bmatrix} 15 & 13 \\ 1 & 6 \end{bmatrix}$ |
| HW | 0.609 | 0.809 | 74.3% | $\begin{bmatrix} 19 & 9 \\ 0 & 7 \end{bmatrix}$ |
| HE | 0.609 | 0.809 | 74.3% | $\begin{bmatrix} 19 & 9 \\ 0 & 7 \end{bmatrix}$ |
| LD-AW | 0.667 | 0.857 | 80.0% | $\begin{bmatrix} 21 & 7 \\ 0 & 7 \end{bmatrix}$ |
| LD-AW Fusion | 0.667 | 0.885 | 82.9% | $\begin{bmatrix} 23 & 5 \\ 1 & 6 \end{bmatrix}$ |

(DAIC-WOZ)

Table 4: Results obtained by [2]

- Hybrid Words (HW) - Considering a heterogenous combination of acoustic and landmark words
- Hybrid Embeddings (HE) - Combining BoAW, BoLW
- Landmark Dependent Acoustic Words (LD - AW) - Acoustic words are only considered within the onset and offset of a landmark

# Literature Organization – Mind Map

- I worked to organize the literature I had read into a sort of mind map from just a linear note-based form

- Could not take a clear picture to add into the ppt, so have embedded the link

Link: https://mm.tt/app/map/3191361121?t=PWbjmax07a

# Tentative Plan

**Plan for next week**
1. Experimenting with data augmentation in line with work by [3]
2. Refining the literature map structure by adding more fine-grained details
3. Exploring and trying to build the architecture framework for my model

**Relevant Links**
1. Overall project plan and timeline: Link
2. Analysis and notes from relevant papers: Link
3. GitHub documenting everyone's presentations and codes: Link

[3] Context-aware Deep Learning for Multi-modal Depression Detection (Lam et al, 2019) | Paper | Code

End