Depression Detection

- Week 7

Name: Sinchana Kumbale

University: Manipal Institute of Technology

Internship Period: Jan – July 2024

Undergraduate 4th year

Duration of the presentation: ~10 minutes

Agenda

1. **Micro Average metrics for the anxiety detection methodology**
2. **Insights from literature review – Paper 1: Data Augmentation Technique**
3. **Insights from literature review – Paper 2**
    1. Understanding the topic division in DAIC-WOZ
    2. Methodology and results
4. **Linguistic Analysis of DAIC-WOZ including the std dev**
5. **Tentative plan for next week**

Micro Average metrics for anxiety detection method

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

Method	Precision	Recall	F1	Accuracy	TP, FP, TN, FN	AUC-ROC
Raw Text	0.59 / 1.0	0.51 / 0.14	0.60 / 0.25	0.74	2, 0, 33, 12	0.65
Pre-processed Text	0.66 / 0.50	0.57 / 0.21	0.56 / 0.30	0.70	3, 3, 30, 11	0.64
Emotional Text	0.68 / 1.0	0.61 / 0.21	0.60 / 0.35	0.76	3, 0, 32, 11	0.68

- Potential Problems caused by very low percentage of depressive samples in dataset (around 20% overall & 30% in the test set).
- Solution is to balance the dataset using Yuxin's upsampling or other techniques

[1] Detecting anxiety from short clips of free form speech (Agarwal et al, 2023) | Paper | Code

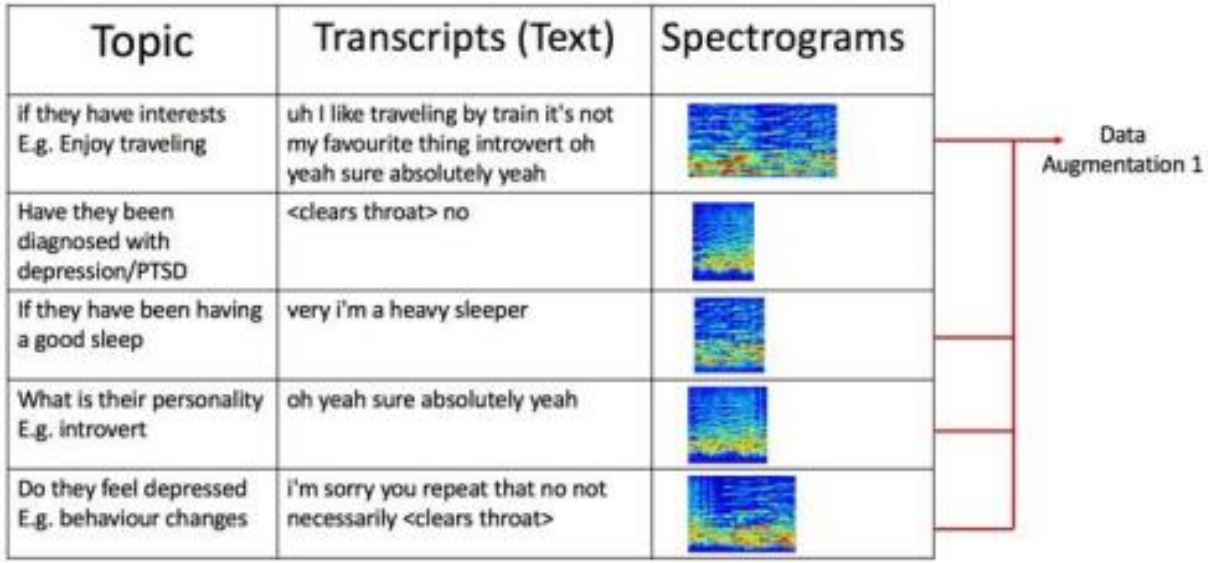
Insights from literature review- Paper 1

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

Fig 1: The topics chosen by [2] for each participant

Topics
If they have interests
About their sleeping habits
How do they feel
Do they feel like a failure
What is their personality
Have they been diagnosed with depression
Their views on parenting

Fig 2: The data augmentation technique [2]



Insights from literature review - Paper 1

Fig 3: The model architecture by [2]

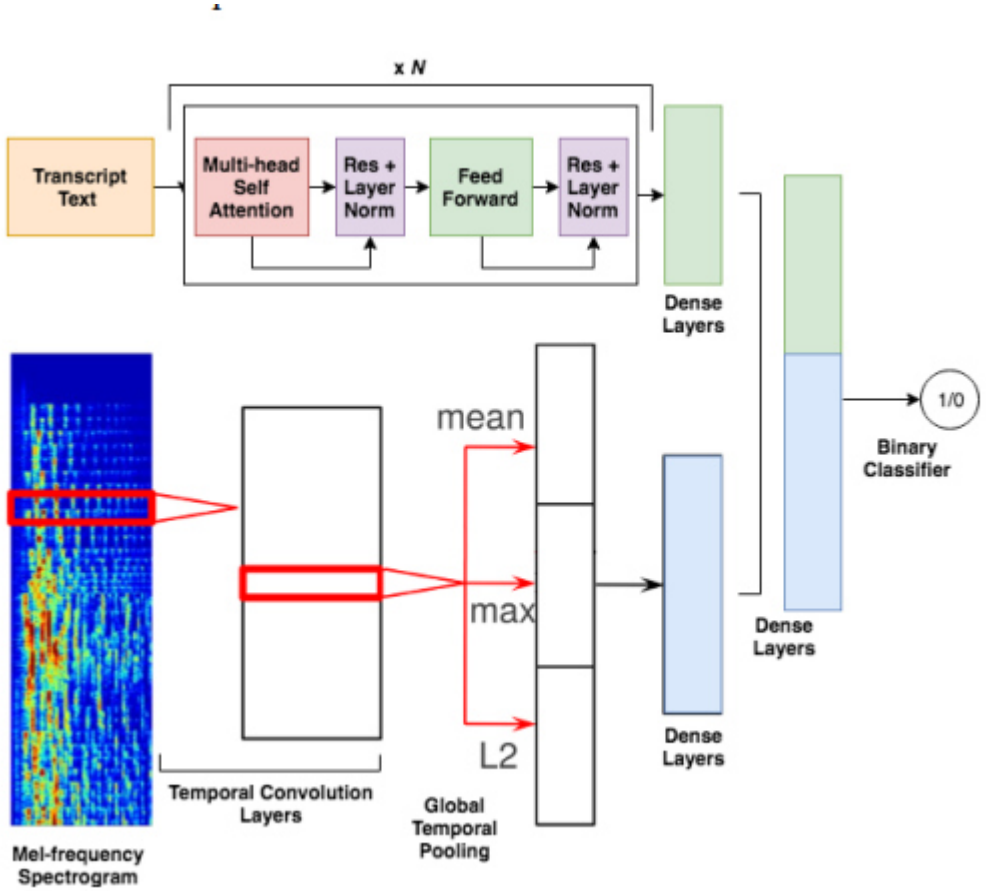


Table 2: The results presented by [2]

Model	Features Type	F1	Precision	Recall
Williamson et al., 2016	Audio Features	0.50	-	-
Williamson et al., 2016	Text	0.76	-	-
Williamson et al., 2016 *	Text	0.84	-	-
Alhanai et al., 2018	Audio Features	0.63	0.71	0.56
Alhanai et al., 2018	Text	0.67	0.57	0.80
Alhanai et al., 2018	Text + Audio Features	0.77	0.71	0.83
CNN-Full	Audio Features	0.56	0.44	**0.78**
CNN-Topic	Audio Features	0.63	0.60	0.67
CNN-Augm	Audio Features	**0.67**	**0.78**	0.58
Trf-Full	Text	0.45	0.37	0.58
Trf-Topic	Text	0.71	0.55	**1.0**
Trf-Augm	Text	**0.78**	**0.82**	0.75
Trf+CNN-Full	Text + Audio Features	0.67	0.60	0.75
Trf+CNN-Topic	Text + Audio Features	0.69	0.64	0.75
Trf+CNN-Augm	Text + Audio Features	**0.87**	**0.91**	**0.83**

Insights from literature review – Paper 2

[3] Topic Modeling Based Multi-modal Depression Detection (Guo et al, 2017) | Paper

Fig 4: The key topics chosen by [3]

76*	change_behavior	have you noticed any changes in your behavior or thoughts lately
77*	depression	have you been diagnosed with depression
78*	easy_sleep	how easy is it for you to get a good night sleep
79*	family_close	how close are you to your family
80*	feeling_lately	how have you been feeling lately
81*	shy_outgoing	do you consider yourself an introvert
82*	ptsd	have you ever been diagnosed with p_t_s_d
83*	therapy_useful	do you feel like therapy is useful

- Observed that overall, in the entire interviews, Ellie initiated 83 topics each having only 1-3 questions
- Found all unique utterances of Ellie, selected those that began a conversation, manually combined the ones which belonged to the same topic
- Around 14 topics cover 80% of the interviews

Insights from literature review – Paper 2

Table 3: The features and dimensions to the SGR model [3]

Feature Name	Dimension
Gender	1
Topic Presence	83
Key Topic	8
LIWC	7719
Formant	1245
COVAREP	18426
AUs	4980
Sum	32462

Table 4: The results presented by [3]

	RMSE			MAE			CC			F1-Score		
	CV	Dev	Test	CV	Dev	Test	CV	Dev	Test	CV	Dev	Test
Basic baseline	5.84	6.57	/	4.81	5.50	/	-0.35	0.00	/	0.00	0.00	/
Challenge baseline	/	7.13	6.97	/	5.88	6.12	/	/	/	/	/	/
Context-unaware baseline	5.55	5.02	/	4.56	4.42	/	0.45	0.69	/	0.58	0.67	/
Proposed method	**3.68**	**3.54**	**4.99**	**2.94**	**2.77**	**3.96**	**0.78**	**0.87**	/	**0.80**	**0.70**	**0.60**

Linguistic Analysis of DAIC-WOZ including the std dev

Table 5: The mean/std dev for linguistic features in the shortened dataset

	pronoun	absolutist	laugh	um	sniffle	sigh	depressive	negative	positive
Non-depressed	19.17/ 12.98	6.31/ 7.27	3.63/ 5.76	18.17/ 16.86	0.30/ 1.98	2.18/ 4.32	1.32/ 3.47	32.58/ 13.05	64.96/ 27.41
depressed	20.55/ 12.29	7.52/ 6.12	5.24/ 7.68	17.50/ 15.70	0.00/ 0.00	2.34/ 4.30	2.61/ 4.23	34.94/ 13.20	49.84/ 16.96

Table 6: The mean/std dev for linguistic features in the complete transcript

	pronoun	absolutist	laugh	um	sniffle	sigh	depressive	negative	positive
Non-depressed	27.00/ 9.24	8.54/ 4.35	6.04/ 6.63	26.82/ 19.74	0.35/ 1.24	2.51/ 3.80	1.56/ 1.62	19.85/ 6.29	52.30/ 13.15
depressed	29.36/9.91	9.27/ 4.29	6.40/ 6.39	25.31/ 18.66	0.25/ 0.72	2.88/ 4.80	2.27/ 2.28	22.75/ 6.04	49.61/ 11.96

[4] What reveals about depression level? The role of multimodal features at the level of interview questions (Shat et al, 2020) | Paper | Code

Value = (count/num_of_words) * 1000

Tentative Plan

Plan for next week

1. Implementing Yuxin's upsampling technique to the anxiety detection code
2. Modifying the shortened dataset with only emotional responses to follow a topic-based structure
3. Compare performance of the different shortened datasets with Yuxin's and the anxiety detection methodology

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

End