

# Paper - 1 Final Report

Name      Nikhil Meena  
Roll No.   210667  
Group      C

## Implementation Details

The codes provided by the author on github is buggy, has lots of missing files that have been used and does not work at all. Additionally, it lacks any instructions on how to extract the dataset to train the model, etc..

So, I have rewritten all the codes from extracting the dataset to training the model and some portion of the code to check the model accuracy.

Additionally, due to size upload limit on github, 3 files namely '*import\_files/vggish/vggish\_model.ckpt*', '*import\_files/zhs.model/encoder.pkl*' and '*import\_files/zhs.model/token\_embedder.pkl*' are uploaded to google drive instead. So, before running the code, the files need to be put to their desired location.

Additional files link : [drive.google.com/...](https://drive.google.com/...)

Github Repo Link : [github.com/Akhilstaar/Multimodal...](https://github.com/Akhilstaar/Multimodal...)

The Model has not been evaluated on DAIC-WOZ dataset because the dataset is not publicly available and is provided on request, I applied for getting access and they provided the dataset in zip format. But, the size of the dataset was too large 100GB total, which required further processing for training.

Also, no significant information has been provided by the author on how to train the model on the DAIC dataset and no code for training on the dataset has been provided.

Additionally, since the data is not used frequently, the download speed were too slow( 100kbps) to download it in reasonable time.

## Results

- **EATD-Corpus Dataset:** This dataset consists of audio and text files from 162 volunteers who received counseling, totaling 518MB. It is publicly accessible at the following link: [EATD-Corpus Link](#).
- For Text Bi-LSTM Model

Metrics	Precision	Recall	F1 Score
<b>My Results</b>	0.78	0.58	0.66
<b>Author's Results</b>	0.65	0.66	0.65

Table 1: Bi-LSTM Model

- For Audio GRU Model

Metrics	Precision	Recall	F1 Score
<b>My Results</b>	0.51	0.66	0.57
<b>Author's Results</b>	0.57	0.78	0.66

Table 2: Audio GRU Model

- For Fuse-Net Model

<b>Metrics</b>	<b>Precision</b>	<b>Recall</b>	<b>F1 Score</b>
<b>My Results</b>	0.65	0.75	0.68
<b>Author's Results</b>	0.62	0.84	0.71

Table 3: Fuse Net Model