Sentimental analysis  
for verifying our first hypo we used sentiment anlasis ,in this we have taken sentiment score and positive and negative word count , this is to check that non depressed person use more Positive word than depressed person and Vice Verca

For this we have used different tools such as textblob,vedar,sentwordnet etc for this sentiment analysis

This tools have predefined library which have sentiment score and we compre each words in transcript with this library and take the overall score and counts

In Dataset 1 each participant has 15 sentneces each and we have performed the analysis for each 15 sentnece and for all sentence at participant level

And in Dataset 2 we have taken the full response of participant and each response of participant and performed the analysis

* Sentiment score
* Avg sent score
* Pos &neg sent score and there averages
* Positive and negative count and there averages
* We have performed normalization for above items
* But we were nt able to distuiguish depressed and non depressed a separate range for this , we were getting overlaps between depressed and non depreseed   
  also by doing this analysis we would get only the world level emotions so we moved to valance and aurosal score

Valence arousal

As we know valence is the level of emotion like low means sad and all and high means happy and all

And also arousal is the level of activness and low means like not interested and all and high means excited and all

And in this also we have performed analysis in both dataset as in sentiment analysis and for this we have used the tool NRC-VAD lexicon for the analysis and it is a predefined lexicon which contain the valence and arousal socre for word and wehave compared the transcripts with words and plotted the following graph

* Aurosal and Valance score and their avg
* Normalised Valance and aurosal for 4 categories and their combinations
* We have used a different thresholds and calculated the EER and at which threshold we would get the best EER we have plotted that
* For Both dataset we have done the analysis

We cannot find any relation or range for depressed vs non depressed all the graphs are overlapped we think that using world level analysis and score we cannot identify the Participant emotion so we have think to consider the sentenece level emotion so we moved to LLMs

LLM

Using LLM we have used two models bert base models and Llama model and we have taken the Sentiment of the responses and the confident score how positive and negative the sentiment is

Using TinyLlama model we get the result and we have manually checked the response and sentiment for most of the sentiment or emotions are correctly taken as compared to the tool we have used

Also we have manually placed the responses in cahtgtp and tried to label the positive and negative emotions like most of them coorectlz labelled to PHQ score

And using Bert base model we have done the analysis but the confident score is close to 1

And also we have tried to do with other Llama models like meta Llama 2 and 3 our computational efficiencz exceeds we cannot run it

So apart from the 2 analysis LLM models are performing the sentimental analysis but we need to look for a better performance for sentimental analysis

The problems arised

Some of the models exceeds computational efficiency

Some of the causes errors

Take too much time to run

So Using LLMs we hope to get some insights