

Firstly, I looked at the datasets and picked *Sentiment Analysis: Emotion in Text*. It is because I worked with sentiment and language analysis on audio-video-text data before, so it was familiar. Secondly, I went for possible techniques to use on this particular dataset and prioritized them based on the timelimit. I wanted to implement classification, clustering, language synthesis, prediction using language property. As per the plan, I implemented the classifier first and then language synthesis. Thirdly, I moved to writing this task-3, which I revised again in the end. At this stage I had to take a long break to attend and finish some of my lab research work. As task-2 needed to put deeper thought into it, I started planning the proposal even before writing it. However, the proposal could have been better if I had more time to make thorough revisions. Especially the prototype building and survey designs could have been more detailed. What went well were – (a) I got to complete my planned techniques in time, (b) got deeper insights of the data (e.g., LIWC analysis) which can be used to strengthen the classifier if needed, (c) The proposal have my primary vision and necessary steps. What did not go so well were – (a) classifiers are overfitting as it can be seen from the training and test accuracy difference, (b) The proposal could have had more reference and clearer descriptions. If I have extra time, I would have – (a) commented out the codes, (b) worked more on regularization of classifier to reduce overfitting, (c) applied other techniques like PCA/Kmeans, (d) worked on the language analysis part to integrate it with the classifiers, (e) revised the proposal multiple times to make it precise and detailed.