**Needs/Risks**

Some of the major preliminary things we need for this project are datasets, reputable and relevant research papers and articles, and a software architecture plan. We have done extensive research and have found all the datasets and papers that we plan on using and can help us in our approach. We also have a robust software architecture plan and data pipeline for our project which has been explained in our technologies and architecture diagrams. In the later stages, we are going to need users for testing and validating our application as well as suggest better UI/UX upgrades to the application. These are some of the primary needs we see arising from our application.

Some of the general concerns of the team regarding this assignment arise from us being successfully able to identify mental illnesses like depression, anxiety, stress, etc. just from a few short conversations with a chatbot. Our hypothesis is that these mental illnesses come along with a certain characteristic tone and language that can be picked up in a conversation with the chatbot. But we are afraid that lacking the non-verbal aspects of conversation such as body language, facial expressions, etc. might negatively impact our model’s accuracy.

We have been able to find some good research papers that enable us to perform novel NLP applications like sentiment analysis, and other speech/text recognition patterns to identify and classify mental illness. But most of these papers use complex word-embeddings and other complex models to accurately perform these NLP applications. We are afraid that we might not be able to accurately replicate the works of these papers while adding our layers of novelty to this algorithm.

With busy schedules for our team members, we are afraid if we have either over-scoped or under-scoped our project. We all are of the opinion that the scope of this proposed project is reasonable, while at the same time is an interesting idea that could use some additional novelty to the algorithm. This is not a major risk/concern per se, but this is something that could arise as a potential problem in the future with our team members’ busy and changing schedules.

Thus, scope creep, unable to implement paper techniques, low accuracy for our algorithm is some of the potential risks for our plan.

**Project Timeline**

As you can infer from the Gannt chart, we have divided our project into 10 weeklong sprints. These phases are logically structured and tasked appropriately to make them doable within a week while making meaningful progress towards the final project.

Here is a quick breakdown of what we plan on doing in the 10 sprints. The Gantt chart provides a more detailed view of the individual tasks assigned to each sprint.

Sprint 1 – Form project teams and explore potential project ideas.

Sprint 2 – Finalize project idea, and find potential datasets and other relevant research papers.

Sprint 3 – Data collection, cleaning, and preprocessing.

Sprint 4 – Creating UI mockups and starting coding out UI screens.

Sprint 5 – Finish working on UI screens and integrate chatbot functionality into the application.

Sprint 6 – Create ML models, NLP applications (LSTM, embeddings, etc.), and integrate backend and UI.

Sprint 7 – Train ML models and get preliminary results for the model.

Sprint 8 – Improve models and result accuracy, complete UI and backend integration, and start conducting user testing and interviews.

Sprint 9 – Prioritize more user interviews and improve the UI and UX features of the application.

Sprint 10 – Work on final touchups to the application and work on the final presentation and report.