Project Progress Report

1) Which tasks have been completed?

   Requirement gathering and planning have been completed – a tech review was conducted to explore the various solutions explored in the current space of analyzing student feedback for understanding areas of improvement for a course. A public GitHub project was also created with a project proposal. Dataset preparation and preprocessing are also in progress and almost completed with general datasets from Kaggle, and synthetic datasets being used for training the Machine Learning (ML) model.

2) Which tasks are pending?

   Developing the ML model, developing a simple frontend for accepting feedback to classify if the general sentiment is positive/negative/neutral. Finally testing and deploying the solution (if possible) along with necessary documentation and presentation for reproducibility.

3) Are you facing any challenges?

   Currently, we cannot find datasets on student feedback closer to UIUC course material. We have found datasets having general feedback and tried creating synthetic datasets. Still, we would like to train our ML algorithms closer to real anonymized data. It would be helpful if there were some data related to the UIUC courses to conduct relevant sentiment analysis.

Task	Hours Required	Comments
Requirement Gathering and Planning	5	Completed – Tech Review conducted to review current problem areas in improving course material by using student feedback and how the latest NLP techniques, large language models and ML algorithms can help resolve some of the issues
Dataset preparation and collection	5	In-Progress – Datasets from Kaggle and other sites have been explored for student feedback for training and testing ML algorithms
Preprocessing	5	In-Progress – Datasets retrieved are being pre-processed to remove stop words, stemming and making the data suitable for effective training
Coding – Classifying feedback based on sentiment analysis using ML algorithms, creating frontend	15	Pending – Using ML model trained and NLP libraries in Python we shall classify if a given feedback is positive/negative. Work will begin on the same this week.
Testing, Deployment, Documentation and Presentation	10	Pending – Will begin from the week starting 27th Nov