

**Instructions:**

Each team must submit a weekly project status form detailing progress, milestones achieved, challenges encountered, and planned next steps. Make sure to include a summary of meetings with the client, summarizing discussions, outcomes, anticipated challenges, and future actions. Also, a detailed project status section captures the accomplishments for the week for technical deliverables and tasks completed. Additionally, please provide any feedback or support that the team may need as the project advances.

Overall Progress:

Provide the overall status of the project, details of discussions, next steps, any challenges and confirm if team members are meeting and contributing.

The project continues to move forward, with significant technical developments this week. We successfully integrated OpenAI Whisper for speech-to-text conversion and RoBERTa for text sentiment analysis. These models were tested on the synthetic dataset to evaluate their performance. Currently, we are focusing on refining the models and integrating them with our data pipeline for audio and text data processing. Our ongoing work includes optimizing the models for key performance metrics such as confidence, precision, and F1 score. Additionally, we are researching deployment strategies to ensure that the pipeline is cloud-ready, enabling scalability for future real-world data processing. The team has remained consistent in meeting deadlines and contributing effectively to the project's goals.

Meeting Summary:

Include the meeting's date and time, summarize key points or outcomes, highlight anticipated challenges, outline the project's next steps, and confirm the attendance of all group members (participants).

Our recent meeting took place on Sunday, starting at 5 PM and concluding at 8:30 PM, with a follow-up session from 10 PM to 1 AM on Monday. During these discussions, the team evaluated the integration of OpenAI Whisper and RoBERTa and analyzed the current performance metrics of the models. We explored various approaches to improve pipeline scalability and deliberated on ways to automate the workflow from data ingestion to output generation. All team members were present and actively participated. The meeting resulted in a clear alignment on finalizing the deployment structure and preparing for real data processing after NDA clearance.



Detailed Project Status:

Include the details of the work accomplished in terms of the technical/project deliverables (GitHub commits) and process (Sprints/Tasks in Microsoft Teams/Planner). Note, the tasks should correspond with each of the roles of your team (Team Charter).

Release Planning: Detail current release with a summary of significant commits in GitHub since last progress report.

The current release includes major updates to the data pipeline, specifically integrating Whisper for speech-to-text and RoBERTa for text analysis. These models have been tested with synthetic data, and we've made significant modifications to enhance the pipeline's scalability and performance. The system is now structured to support larger datasets, and once actual data is available, we will initiate final testing to validate the system's effectiveness in a production environment.

Sprint Planning: Detail current sprint with a summary of task progress (new tasks, completed tasks in Microsoft Teams) since last progress report.

In **Sprint 2**, our efforts were centered around preliminary testing and development using synthetic data while awaiting NDA clearance for real data access. We conducted thorough research on open-source models, leading to the selection of OpenAI Whisper for speech-to-text and RoBERTa for sentiment analysis. Additionally, we finalized the process flow for data ingestion, model training, and evaluation, ensuring compliance with project requirements.

In **Sprint 3**, we have transitioned into full implementation and comprehensive evaluation of these models. The main focus is on enhancing their performance in terms of accuracy, precision, and scalability, using metrics like confidence and F1 score. Simultaneously, we are building a cloud-ready data pipeline that automates the entire workflow from ingestion to output generation. This will allow us to efficiently process real-world data once NDA clearance is obtained and prepare the system for future scalability.

Feedback/Support:



Please list any questions or concerns for the professor or TA, and specify any resources or guidance needed for the project.

We do not have any questions for the Professor.