Meeting Notes

Week 1: Defining the Use Case and Agreeing on What Model to Use

- **Date and Time:** September 29th at 1:00 PM
- **Attendees:** Tin, Somesh, Rishabh
- **Note Taker:** Tin
- **Meeting Notes:**
    - Discussed three use cases for the Virtual Assistants (VAs): Personal Fitness Advisor, Opera Cloud Digital Assistant, Intelligent Email Parsing and Information Retrieval with RAG.
    - Agreed to use LLaMA and Gemma models for fine-tuning due to their performance and adaptability.
    - Established weekly milestones for the project timeline.

Week 2: Assemble Document Corpus

- **Date and Time:** October 6th at 1:00 PM
- **Attendees:** Tin, Somesh, Rishabh
- **Note Taker:** Somesh
- **Meeting Notes:**
    - Identified document corpus requirements for each use case.
    - Sources selected:
        - Tin: Fitness data from public fitness and health resources.
        - Rishabh: Opera Cloud documentation from vendor-provided APIs and manuals.
        - Somesh: Email parsing and retrieval documents from public datasets and proprietary resources.
    - Assigned tasks for assembling and preprocessing the corpus.

Week 3: Set Up the LLMs (LLaMA and Gemma)

- **Date and Time:** October 13 at 1:00 PM
- **Attendees:** Tin, Somesh, Rishabh
- **Note Taker:** Rishabh
- **Meeting Notes:**
    - Set up environments for LLaMA and Gemma models.
    - Verified GPU availability and configured training pipelines.
    - Addressed initial challenges with model compatibility and dependencies.
    - Defined a tentative timeline for completing model setup.

Week 4: Implement VA Using Each Model in RAG Pipeline

- **Date and Time:** October 20 at 1:00 PM
- **Attendees:** Tin, Somesh, Rishabh
- **Note Taker:** Tin

- **Meeting Notes:**
    - Implemented Retrieval-Augmented Generation (RAG) pipeline for each VA.
    - Discussed architecture and modularity to adapt the pipeline to each model.
    - Verified document retrieval performance with initial queries.
    - Highlighted improvements needed for retrieval speed and relevance.

Week 5: Evaluate the Accuracy of the VA

- **Date and Time:** October 27 at 1:00 PM
- **Attendees:** Tin, Somesh, Rishabh
- **Note Taker:** Somesh
- **Meeting Notes:**
    - Conducted accuracy evaluations for all VAs.
    - Metrics used: relevance, precision, recall, and user satisfaction.
    - Observed strengths:
        - LLaMA: Strong in general knowledge and adaptability.
        - Gemma: Exceptional in specialized tasks like Opera Cloud assistance.
    - Identified weaknesses: LLaMA struggled with highly domain-specific queries; Gemma required additional fine-tuning for email parsing.

Week 6: Improve the Performance of the VA

- **Date and Time:** November 10th at 1:00 PM
- **Attendees:** Tin, Somesh, Rishabh
- **Note Taker:** Rishabh
- **Meeting Notes:**
    - Implemented improvements based on Week 5 evaluations:
        - Enhanced fine-tuning of Gemma for email parsing.
        - Optimized RAG pipelines for better retrieval speed.
        - Added domain-specific data to LLaMA for improved context understanding.
    - Tested updates and noted significant accuracy improvements.

Week 7: Present Results

- **Date and Time:** November 17th at 1:00 PM
- **Attendees:** Tin, Somesh, Rishabh
- **Note Taker:** Tin
- **Meeting Notes:**
    - Presented the final VAs and their respective use cases.
    - Demonstrated live queries and system performance.
    - Received feedback from team members and documented observations for potential future work.
    - Finalized the project documentation and results to class.

Peer Review

From Rishabh

To Somesh Bagadiya - Somesh Suggested different models that we could use. He was well-informed with several different models that are being used in the market. He was mostly active in giving prompt responses to any blocker or doubts

To Tin Pho - Tin is good with scraping web data and understanding the structure of web pages. He helped write good parsers to scrape data. He was also available to give peer review of my code and how I could enhance it

From Somesh

To Tin Pho - Tin has been an exceptional teammate, always ready to assist with valuable insights and support. Their expertise in web scraping and data preprocessing has been incredibly helpful, and their ability to troubleshoot issues and suggest solutions has greatly benefited the team. Tin is approachable, collaborative, and a reliable contributor who elevates the team's overall performance.

To Rishabh Jain - Rishabh has been instrumental in optimizing the RAG pipeline, offering insightful suggestions that improved performance and accuracy. His deep understanding of RAG parameters and ability to share practical tips have been a significant asset to the team. Rishabh's collaborative approach and technical acumen are commendable.

From Tin Pho

To Somesh Bagadiya - Somesh helps me a lot with implementing the models. He guides me through the process, explaining complex concepts in a way that's easy to understand. Whether it's setting up the model's structure, optimizing its performance, or troubleshooting issues, he is always there to offer support. His knowledge and patience make the entire implementation process much smoother and more efficient.

To Rishabh Jain - Rishabh has helped me by sharing insights on RAG optimization and suggesting evaluation metrics. He clearly explained the concepts of RAG to me when I did not fully understand it. He also shared results from his training experiments to benefit the whole team.