# Team Contribution Report

Group – 2

University of Missouri – Kansas City COMP\_SCI5588: Data Science Capstone Yugyung Lee

April 3rd, 2024

# Individual Contribution Report

**Deepak Ayyasamy: Machine Learning Development**

**Ingestion Module Development**:

* Developed ingest.py module responsible for loading documents from specified directories and splitting them into appropriate formats.
* Implemented multithreading and multiprocessing techniques for efficient document loading and processing.
* Incorporated error handling mechanisms to manage exceptions during the document loading process.
* Created a logging system (file\_ingest.log) to track document loading progress and any encountered errors.

# Local GPT Module Development:

* Developed run\_localGPT.py module responsible for running retrieval question-answer tasks using Language Chain's RetrievalQA pipeline.
* Implemented model loading functions to load pre-trained models based on specified configurations.
* Integrated embeddings and vector stores for efficient text representation and retrieval.
* Leveraged Mistral 7B quantized version as LLM model and “instructor-large” as embedding model.
* Implemented callbacks for streaming responses and managing output verbosity.
* Provided options for users to specify device type, show/hide source documents, use history, and select model types.

# Code Refinement and Documentation:

* Refactored codebase for improved readability and maintainability.
* Added comprehensive inline documentation to explain module functionalities and code logic.
* Conducted thorough testing and debugging to ensure code correctness and reliability.
* Incorporated user-friendly CLI interfaces using the Click library for easy parameter configuration.

# Challenges Faced:

* **Concurrency Management**: Implementing efficient concurrency management strategies for document loading and processing required careful consideration to avoid race conditions and deadlocks.
* Addressed issues related to thread and process synchronization to ensure smooth execution.
* **Model Loading and Configuration:** Ensuring compatibility and proper configuration of pre-trained models posed challenges, especially in managing different model types and device types.
* Implemented robust error handling mechanisms to handle model loading failures and configuration errors gracefully.
* **Optimization and Performance Tuning:** Optimizing document loading and processing performance while minimizing resource utilization was a significant challenge.
* Explored various optimization techniques and fine-tuned parameters to achieve a balance between performance and resource efficiency.
* **Individual Member Contributions in Percentage:**

Deepak Ayyasamy : 33.3%