**ADTA 5770: Generative AI with LLM**

**Semester Project – Individual Final Report**

Student Name: Biniam Abebe

Group: 8

1. **Overview**

The semester project aims to develop a knowledge-based Question-Answer and Search System. The project is done using the Cloud Integrated Development Environment (IDE) System (CIDES) provided by Google Cloud Platform (GCP) Vertex AI services.

Our group developed "Financial Intelligence," a Generative AI-powered Q&A system designed for corporate financial analysis. This system enables employees to search proprietary financial documents efficiently, ask questions, and receive accurate answers, ultimately streamlining workflows and improving decision-making processes.

1. **Semester Project: Submitted Individual Portfolio**

Below is a list of all documents and materials I have submitted as part of my individual portfolio:

1. Final Project Proposal - "Financial Intelligence: A Generative AI-Powered Q&A System for Corporate Financial Analysis"
2. Part III: Knowledge Base - Collection of 100 PDF documents related to corporate finance (uploaded to assigned OneDrive folder)
3. Part IV: System Planning, Requirements, and Data - Documentation of business organization selection, business and technical requirements, and GCP setup
4. Part V: System Analysis - Comprehensive analysis document including problem statement, feasibility analysis, and project management plan
5. Part VI: System Design - High-level and detailed system design documentation
6. System Development Documentation - Python code implementation using LangChain and Vertex AI
7. System Testing Results - Performance evaluation and user feedback
8. Individual Final Report - This document
9. **Semester Project: Individual Contributions**

Throughout this semester project, I made the following contributions to our group's development of the Financial Intelligence Q&A system:

**Project Planning and Requirements Phase**

* Participated actively in brainstorming sessions to define our domain expertise in corporate finance
* Led the technical components section of our project proposal, researching and documenting implementation approaches

**Knowledge Base Development**

* Collected and curated all the 100 PDF documents for our knowledge base
* Specialized in gathering quarterly financial reports from Fortune 500 companies and regulatory guidance documents from the SEC
* Developed a categorization system for our documents to facilitate efficient retrieval and context management

**System Setup and Configuration**

* Set up the Google Cloud Platform environment, including API configurations and service account permissions
* Created and configured the primary storage buckets (adta5770-docs-folder98) for document management

**System Analysis and Design**

* Authored the technical feasibility analysis section of our system analysis document
* Designed the vector embedding component of our system architecture, focusing on optimizing financial terminology representation
* Created detailed data flow diagrams illustrating the relationship between system components

**System Development**

* Implemented the document processing pipeline using Python, LangChain, and PyPDF2
* Developed the embedding generation module that converts financial documents into vector representations
* Created an efficient chunking algorithm that preserves the context of numerical data in financial statements
* Integrated the Retrieval Augmented Generation (RAG) components to ensure accurate responses based on source documents

**Testing and Refinement**

* Designed a comprehensive test suite with 20+ test cases covering various financial query types
* Implemented query refinement algorithms based on test results to improve response quality for complex financial questions

1. **Semester Project: Teamwork Experience**

**Positive Experiences**

* Our team established a structured approach to collaboration from the beginning, with clear role assignments and regular check-ins. This organization helped us stay on track throughout the semester and ensured everyone understood their responsibilities.
* The diverse skills within our team complemented each other well. Biniam and Joshua brought strong domain expertise in finance, while Srilekha and Nithin contributed advanced technical knowledge of AI systems and cloud platforms.
* We maintained open communication through multiple channels (group chat, video meetings, shared documents), which facilitated efficient information sharing and problem-solving. Our weekly Zoom meetings were particularly effective for aligning on key decisions.
* All team members consistently demonstrated commitment to the project by meeting deadlines, producing high-quality work, and offering assistance to others when needed.
* When we encountered challenges with our vector embedding approach, the team came together for an extended troubleshooting session where everyone contributed ideas until we found an optimal solution.
* The collaborative environment fostered innovation, with everyone feeling comfortable to share ideas and constructively critique approaches to reach the best outcomes.

**Areas for Improvement**

* While our teamwork was excellent overall, we could potentially benefit from more documentation of our development process to facilitate knowledge sharing and future maintenance.
* We could establish more regular technical demos throughout the development process to ensure everyone maintains a comprehensive understanding of all system components.
* Exploring additional collaboration tools specifically designed for software development might further enhance our workflow efficiency in future projects.

1. **Semester Project: Group Member’s Teamwork Peer-Evaluation**

**Srilekha Aduvala: Excellence** Srilekha's technical expertise was invaluable to our team. She took the lead in setting up our GCP environment and configuring the vector database infrastructure. Her implementation of the RAG architecture was exceptionally well-documented and efficient. Srilekha consistently attended all meetings, contributed thoughtful ideas, and was always willing to help other team members understand complex technical concepts. She was particularly effective in translating technical requirements into practical implementation steps, which helped keep our development process on track.

**Nithin Marpu: Excellence** Nithin demonstrated outstanding commitment to the project by consistently delivering high-quality work. His contributions to the query processing module were exceptional, incorporating sophisticated financial domain knowledge into the system's ability to understand user requests. Nithin was always responsive to communication and provided detailed updates on his progress. When we faced challenges with our embedding model performance, Nithin researched and implemented solutions that significantly improved our system's accuracy. His ability to balance technical implementation with user experience considerations enhanced the overall quality of our final product.

**Joshua Terrazas: Excellence** Joshua's contributions to the project were exemplary. He developed an excellent high-level system design that formed the foundation of our implementation. Joshua took responsibility for the response generation component, ensuring that our system provided accurate, well-formatted financial information. His attention to detail during the testing phase helped identify and address several potential issues before they affected system performance. Joshua was consistently collaborative, sharing resources and offering assistance to team members whenever needed. His positive attitude and problem-solving approach were valuable assets to our team dynamics.