**University of North Texas**

**ADTA 5760-501**

# **Natural Language Processing with Artificial Neural Networks**

**Financial Intelligence: A Generative AI-Powered Q&A System for Corporate Financial Analysis**

**Group Members** – Biniam Abebe, Srilekha Aduvala ,Nithin Marpu, Joshua Terrazas

**Abstract**

This proposal outlines an Enterprise-level Generative AI Q&A Search system designed to facilitate valuable corporate financial statement analysis for Corporate Financial Statements. The second system will serve as a powerful tool for the organization’s employees to conduct precise content searches and get accurate answers about financial data included in the company’s proprietary documents. Using Google Cloud Platform's Vertex AI services and generative AI technologies, its services will help analyze quarterly capital statements for 100 public companies, delivering timely intelligence to improve valuable financial insights.

**Field of Domain Expertise: Corporate Finance**

We have identified our area of expertise as corporate finance and chosen financial statement analysis. Modern business organizations and their processes require fast, scalable solutions to process and interpret complex financial documented data. Our AI model will be trained on the Q4 financial statements of 100 main public companies, ensuring the responses align with real-world reporting standards.

**Technical Environment**

1. Development Platform
   * Google Cloud Platform (GCP): Vertex AI services
   * Google Collaboratory (Colab) Pro account for team development
   * Python as the primary programming language
   * Utilization of the latestI/ML libraries, including TensorFlow
2. Core Technologies
   * Retrieval Augmented Generation (RAG) for enhanced search accuracy
   * Sentence Transformers for semantic similarity matching
   * Gemini 2.0 flash for real-time financials
   * Fine-tuned large language models (LLMs)
3. Infrastructure Components
   * Cloud storage for document management
   * Vector embeddings generation services
   * Vector database management systems for high performance search retrieval
   * Advanced vector search technologies

**Project Implementation Plan**

1. Knowledge Base Development
   * Collection and indexing of 100 Q4 financial statements
   * Document preprocessing and standardization
   * Vector embedding generation
   * Database structure optimization to handle large-scale financial data queries
2. System Architecture
   * Integration with GCP Vertex AI services
   * Implementation of RAG architecture
   * Scalability considerations, including potential expansion to additional quarters
   * Development of search and retrieval mechanisms
   * Query processing and response generation

**Expected Deliverables**

1. Functional System Components
   * End-to-end document processing pipeline
   * Vector database implementation
   * Advanced natural language query processing system
   * Response generation module with contextual financial insights
   * User-friendly interface for financial analysts and corporate decision-makers
2. Performance Metrics
   * Query response accuracy
   * Processing speed metrics
   * System reliability measures
   * User satisfaction metrics
   * Compliance with financial reporting regulations and standards

**Conclusion**

This project seeks to pioneer financial analysis by integrating state-of-the-art Generative AI capabilities with deep financial domain knowledge. With data up to October 2023, the system will deliver recognized insights into these areas that will be useful for economic experts and decision-makers. With the right execution and extensive testing, we will build a powerful tool to elevate our entity's financial analysis.