

FinSights

Name: Arnab Chakraborty's

Course: Prompt Engineering- leverages SEC filings data to inform trading strategies.

Date: July 15, 2024



Introduction

Project Focus: Developing an application to analyze Financial Reports

This project aims to create a tool that processes SEC filings, including Earnings Calls 10-K, 10-Q and management commentary filed by companies to provide a transparent view of their financial performance.

Objective: Inform trading strategies

The application provides financial reports to help investors and traders to identify and analyze stocks, calculate risk and help in the creation of data-driven investing strategies.

Project Overview

Objectives

Leverage SEC filings data using RAG to generate trading insights. Improve data retrieval accuracy and enhance insight quality.

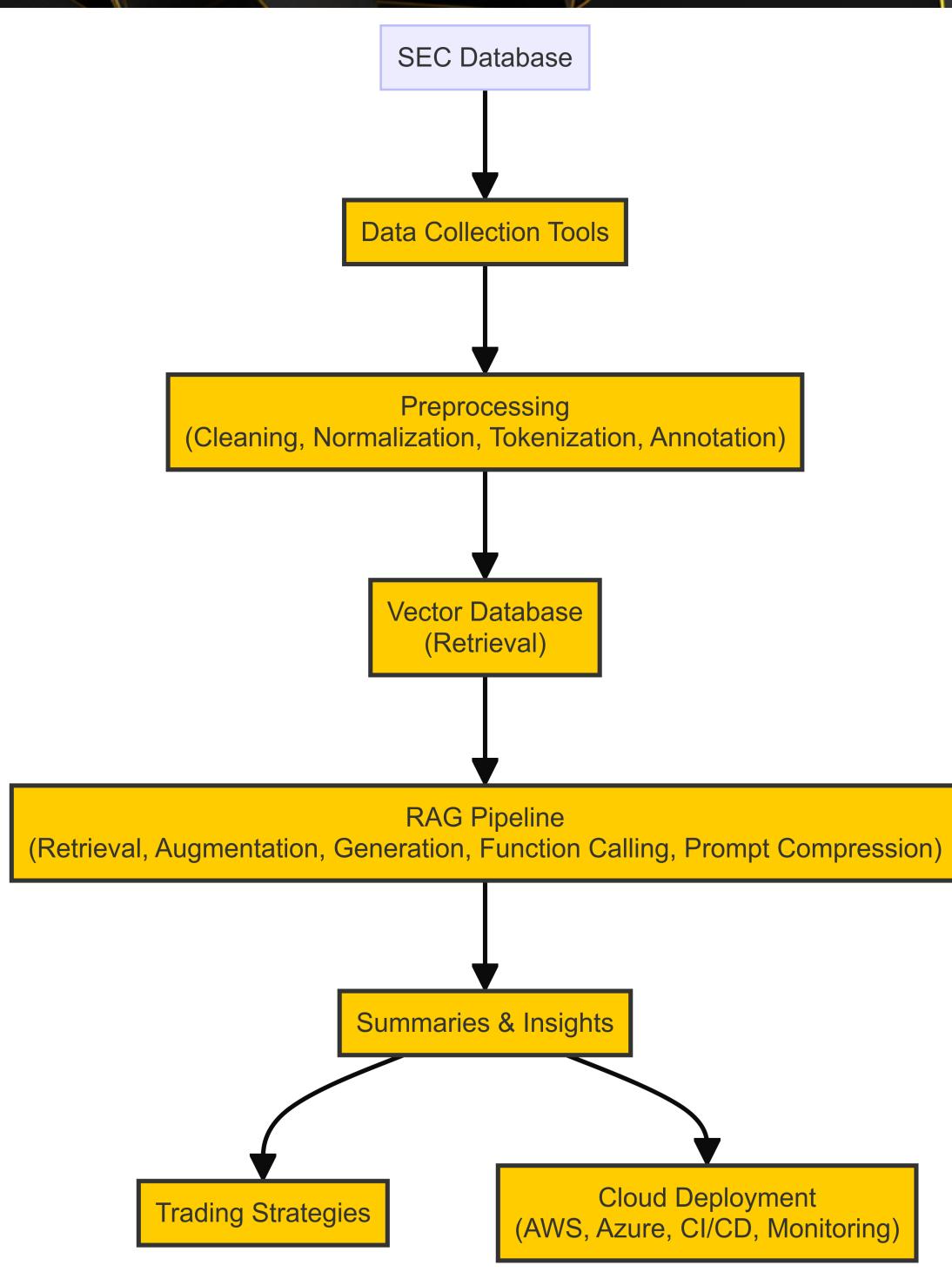
Relevance

Demonstrates practical application of course techniques. Provides efficient analysis tool for financial professionals.

Scope

Analyze SEC filings including 10-Ks, 10-Qs, and earnings transcripts. Implement RAG pipeline with function calling.

Project Architecture



- 1** **Data Collection**
Gather SEC filings from public sources.
- 2** **Preprocessing**
Clean and structure data for analysis.
- 3** **RAG Pipeline**
Integrate retrieval with language models.
- 4** **Output**
Generate insights and trading strategies.

Data Collection and Preprocessing

1. Source

SEC's EDGAR database,
AlphaVantage/ Yahoo Finance
containing regulatory documents.

2. Collection

Use web scraping tools and APIs
to gather and process documents.

3. Preprocessing

Clean, normalize, tokenize, and
annotate data for analysis.



RAG Pipeline Implementation

1

Retrieval

Use vector databases to search and retrieve relevant SEC filing sections.

2

Augmentation

Integrate the Company financial and earning reports to enrich context for analysis.

3

Generation

Employ language models to generate functions/ json output, analyze insights - trends/valuations from retrieved data.

4

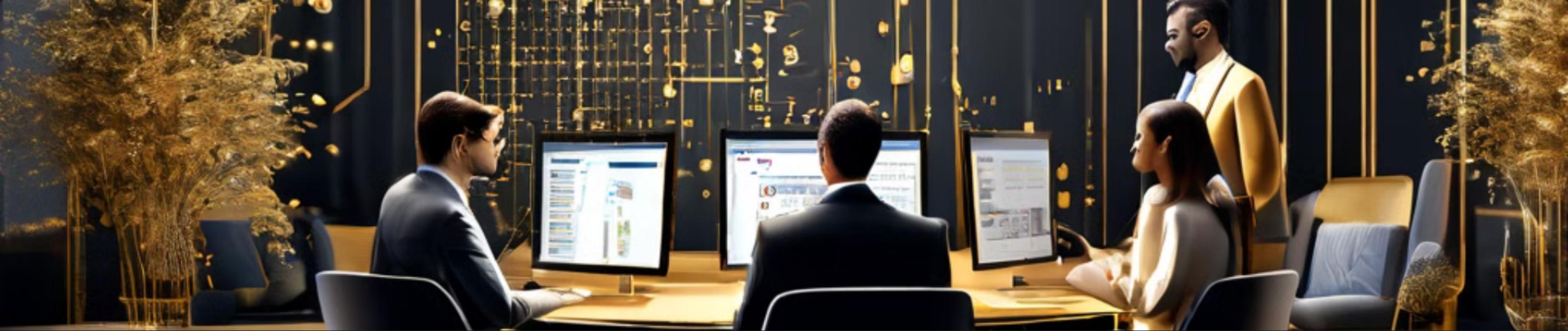
Function Calling

Extract real time specific financial metrics and details from filings.

Performance Metrics

Metric	Method
Faithfulness	Use LLM to evaluate the response from the provided context
Latency	Measure response time
Relevance	Use a LLM to measure the relevance of response & context with Question





Improvement Strategies



Data Augmentation

Enrich training data with additional financial documents (Fund Returns).



Model Tuning

Fine-tune LLM on financial domain (Capture Abbreviation).



Algorithm Optimization

Enhance retrieval using other similarity metrics and reranking.

Deployment and Future Work

Deployment Steps

Deploy the frontend and backend with LLM hosted

User Testing

Release beta version and integrate for RLHF to tune the models.

Future Extensions

Expand to include other financial reports



Conclusion



Key Takeaways

Efficient analysis of Financial documents provides advantages in developing trading strategies.

Impact

RAG with function calling enhances accuracy and relevance of generated insights.