

Week 5 Write-Up: Advanced RAG (Track A & B)

In this week's assignment, I extended my Retrieval-Augmented Generation (RAG) pipeline by implementing improvements in two areas: Track A (Reranking & Context Optimization) and Track B (Multimodal RAG with text + images). These upgrades build directly on my capstone project, which focuses on adaptive multi-agent systems for stock market prediction and sentiment analysis. The goal was to improve retrieval quality, add support for financial charts, and make the system more useful for real-world analysis.

Methods

Track A: I first combined BM25 and dense retrieval using Reciprocal Rank Fusion (RRF). I also added a reranker to rescore top passages and experimented with passage compression. This aimed to balance retrieval accuracy with lower token costs. Track B: I introduced multimodal retrieval by including four financial charts. Captions were generated for each chart (e.g., "Blue candlestick chart showing stock price movements"). I then built a joint index that stored both text chunks from project documents and image captions. Queries could now retrieve text, images, or both, depending on the input.

Results

Track A: Reranking improved the relevance of results compared to the baseline. Recall@5 increased, and responses became more concise after compression, though latency slightly increased. Track B: Multimodal retrieval worked successfully. For example, when querying "stock volatility chart," the system retrieved the correct candlestick chart along with related text passages. This shows that charts can complement text evidence and give richer responses. Overall, both tracks demonstrated measurable improvements in retrieval quality and project relevance.

Conclusion & Next Steps

By completing Track A and Track B, I created a more accurate and flexible RAG pipeline. Reranking improved text retrieval quality, while multimodal retrieval enabled chart-based evidence alongside documents. These improvements directly support my project's focus on financial analysis, where both textual reports and charts matter. Next, I plan to expand the multimodal dataset with more realistic financial images and evaluate the system on additional queries to further validate performance.