

[Finance Report Generation with RAG]

Kiha Kwon, Ina Jung, and Misuk Kim
School of Data Science, Hanyang University

Research Motivations

Analysts must review vast amounts of financial news and reports, yet many LLM-based systems still struggle to retrieve accurate, up-to-date information.

Despite abundant structured and unstructured market data, current approaches often fail to integrate them effectively for reliable report generation.

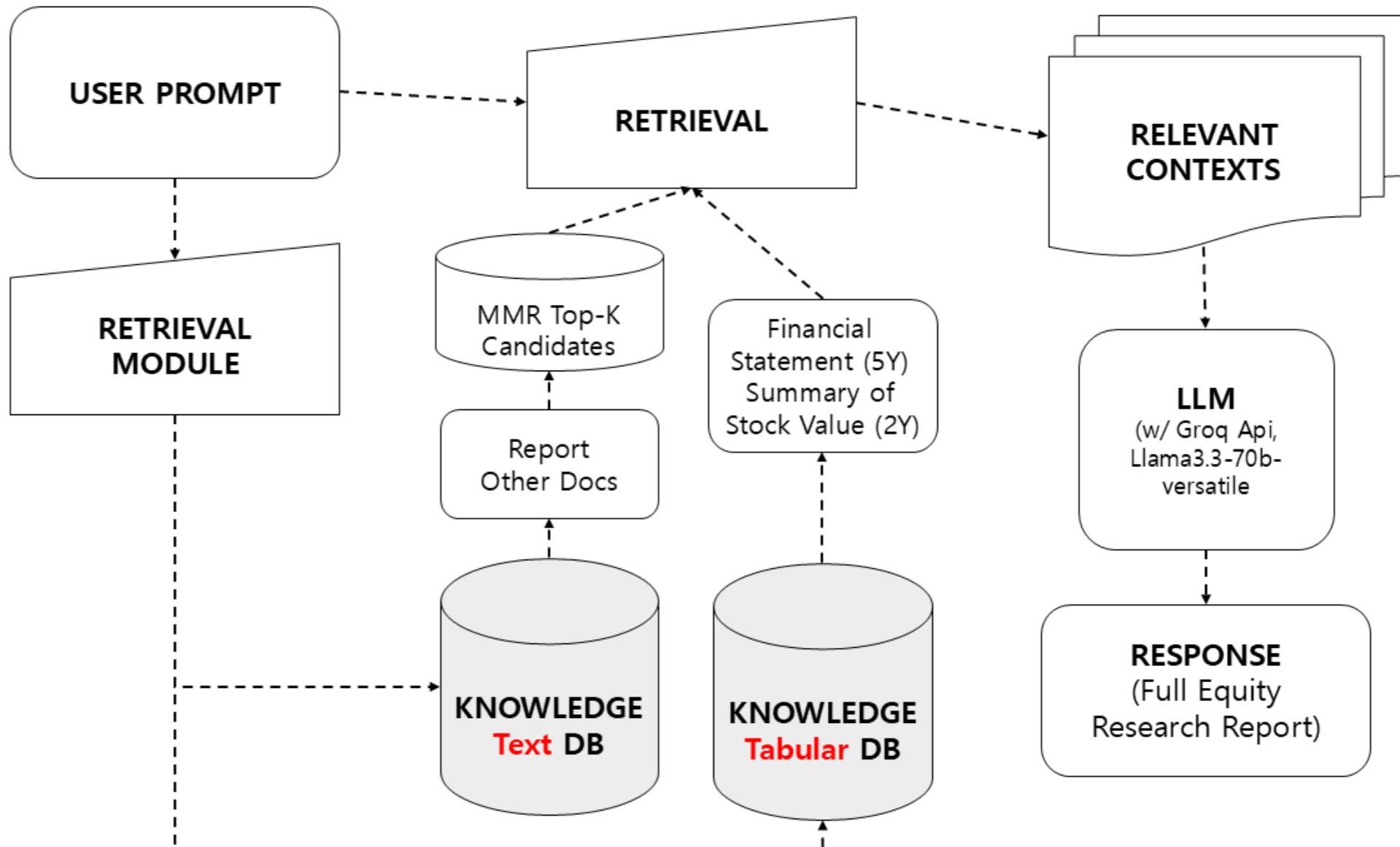
To overcome these gaps, we build a retrieval-augmented system that produces concise, analyst-style financial reports.

Objectives

Processing Input Data: Address the lack of consistency across financial text formats and improve early-stage retrieval accuracy.

Handling Output Quality: Overcome issues where the model produced irrelevant or repetitive responses, even with more data.

Model architecture



Implementation Keywords

- Ingestion:** HTML cleaning, document labeling (name, ticker, keywords...)
- Chunking:** Recursive splitter, firm & ticker extraction
- Embedding/Indexing:** ko-SRoBERTa embeddings, FAISS index
- Hybrid Retrieval:** BM25 + dense search, RRF fusion, CE rerank, MMR
- LLM Generation:** Structured prompt, financial data injection, Groq LLM

Results

삼성전자, 2025년 2분기 실적 개선세 지속 전망

재무 요약

연도	매출액	영업이익	영업이익률
2024	300,870,903,000,000	32,725,961,000,000	10.88
2023	258,935,494,000,000	6,566,976,000,000	2.54
2022	302,231,360,000,000	43,376,630,000,000	14.35
2021	279,604,799,000,000	51,633,856,000,000	18.47
2020	236,806,988,000,000	35,993,876,000,000	15.20

주요 주가 통계

지표	값
max_price	86,207
min_price	49,269
avg_price	66,003
std.dev	9,140
avg_volume	19,039,583

최근 월별 종가

월	종가
2025-02	54,178
2025-03	57,800
2025-04	55,500
2025-05	56,200
2025-06	59,800
2025-07	70,600

요약

최근 삼성전자의 실적은 DS 사업부와 MX의 성과가 주요 원인이다. DS 사업부의 경우, DRAM 출하와 가격이 예상치를 상회하였으며, Legacy 플랫폼도 예상치를 상회하였다. 그러나, HBM은 상대적으로 부진한 것으로 추정된다. MX의 경우, 모바일 성수기 영향으로 지난 해 수준의 영업 성과를 보인 것으로 추정된다. 또한, 월/달러 효과도 실적 개선에 기여한 것으로 추정된다. 삼성전자의 가전제품 판매도 인공지능(AI) 사용 경험 확산을 위한 총력전을 펼치고 있어, 가전제품 판매에도 긍정적인 영향을 미칠 것으로 기대된다. 특히, 동남아시아 지역에서 AI 기능을 갖춘 제품 구매 의향이 높다는 결과가 나왔으며, 이는 삼성전자의 가전제품 판매에 긍정적인 영향을 미칠 것이다.

분석

삼성전자의 최근 실적은 DS 사업부와 MX의 성과가 주요 원인이다. DS 사업부의 경우, DRAM 출하와 가격이 예상치를 상회하였으며, Legacy 플랫폼도 예상치를 상회하였다. 그러나, HBM은 상대적으로 부진한 것으로 추정된다. MX의 경우, 모바일 성수기 영향으로 지난 해 수준의 영업 성과를 보인 것으로 추정된다. 또한, 월/달러 효과도 실적 개선에 기여한 것으로 추정된다. 삼성전자의 가전제품 판매도 인공지능(AI) 사용 경험 확산을 위한 총력전을 펼치고 있어, 가전제품 판매에도 긍정적인 영향을 미칠 것으로 기대된다. 특히, 동남아시아 지역에서 AI 기능을 갖춘 제품 구매 의향이 높다는 결과가 나왔으며, 이는 삼성전자의 가전제품 판매에 긍정적인 영향을 미칠 것이다.

Future Works

- Integrate real-time news & reports for live, auto-updated report generation.
- Improve numeric grounding and expand to multi-company/sector comparisons.