project report fintech.doc

Vishwakarma Group of Institutions

Document Details

Submission ID

trn:oid:::3618:93168018

Submission Date

Apr 27, 2025, 10:02 PM GMT+5:30

Download Date

Apr 27, 2025, 10:03 PM GMT+5:30

File Name

project report fintech.doc

File Size

119.5 KB

13 Pages

3,564 Words

21,801 Characters

10% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography

Match Groups

26 Not Cited or Quoted 10%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

7% Internet sources
5% Publications
8% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **26** Not Cited or Quoted 10%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

7% Internet sources
5% Publications
8% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Submitted works	ICTS on 2024-08-30	2%
2	Internet	arxiv.org	<1%
3	Internet	www.coursehero.com	<1%
4	Internet	curin.chitkara.edu.in	<1%
5	Submitted works	Manipal University on 2023-12-17	<1%
6	Submitted works	University of Stirling on 2024-08-06	<1%
7	Internet	diva-portal.org	<1%
8	Internet	vocal.media	<1%
9	Internet	research.ijcaonline.org	<1%
10	Submitted works	srmap on 2025-04-26	<1%

11 Internet

www.mdpi.com <1%

12 Internet

www.version1.com <1%

13 Publication

Oduba, Ayomide Elijah. "Enhancing E-Commerce Dataset Recommendations Usin... <1%

14 Submitted works

Queen Mary and Westfield College on 2025-03-27 <1%

15 Publication

Sunil Arya. "An optimal algorithm for approximate nearest neighbor searching fi... <1%

16 Submitted works

University of Hertfordshire on 2024-09-02 <1%

17 Submitted works

University of Hong Kong on 2005-05-10 <1%

18 Internet

etd.aau.edu.et <1%

19 Submitted works

srmap on 2025-04-21 <1%

20 Internet

www.careers360.com <1%

21 Internet

www.grin.com <1%

22 Submitted works

Capella University on 2024-04-26 <1%

23 Submitted works

Otto-von-Guericke-Universität Magdeburg on 2025-03-09 <1%

A PRELIMENERY REPORT ON

Enhancing Financial Advisory Systems through Retrieval-Augmented Generation and LLMs

SUBMITTED TO THE VISHWAKARMA INSTITUTE OF INFORMATION TECHNOLOGY, PUNE
IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

OF

BACHELOR OF TECHNOLOGY (COMPUTER SCIENCE &ENGINEERING (AIML))

SUBMITTED BY

STUDENT NAME	Exam Seat No. :
TANISHK SINGH	22211518
MITHILESH GADO	22210486
SAANVI BHAT	22210135
VIPUL KHANDELWAL	22211149

DEPARTMENT OF COMPUTER ENGINEERING

BRACT'S
VISHWAKARMA INSTITUTE OF INFORMATION TECHNOLOGY

SURVEY NO. 3/4, KONDHWA (BUDRUK), PUNE – 411048, MAHARASHTRA (INDIA).

Enhancing Financial Advisory Systems through Retrieval-Augmented Generation and LLMs

01. Introduction

1.1 Overview

Financial advisory systems have come a long way with advances in technology. Conventional financial advisory services tend to be costly and require human advisors, rendering them inaccessible to most people. This project creates an intelligent financial advisory chatbot that relies on Retrieval-Augmented Generation (RAG) and Large Language Models (LLMs) to offer personalized financial advice. The system provides guidance on three principal financial activities: loans, insurance, and investments, rendering expert-level financial planning accessible to all.

1.2 Motivation

The impetus for this project comes from some important observations regarding the present state of finances:

- Financial illiteracy persists despite ready access to information online
- Professional financial advisors are too expensive for many people to pay
- Current robo-advisors employ static models that do not respond to dynamic market changes
- There is a demand for individualized financial advice that takes individual situations into account
- Current advisory systems don't integrate real-time data and context awareness

These factors leave a large gap in financial services that can be filled by using advanced AI technologies.

1.3 Problem Definition and Objectives

Problem Definition: How can we develop an AI-powered financial advisory system that provides personalized, accurate, and accessible financial guidance by combining Retrieval-Augmented Generation (RAG) with Large Language Models while ensuring ethical compliance and user trust?

Objectives:

1. Create domain-specific advisory modules for loans, insurance, and investments
2. Integrate real-time financial data through APIs and web scraping
3. Implement RAG architecture to retrieve relevant financial information
4. Leverage Google Gemini Pro to generate personalized financial recommendations
5. Ensure transparency, explainability, and regulatory compliance
6. Evaluate the system's performance using standard NLP metrics

1.4 Project Scope & Limitations

Scope:

- Development of a chatbot interface for financial advisory services
- Implementation of three specialized advisory modules (loan, insurance, investment)
- Integration of real-time financial data sources
- Natural language processing for user queries
- Retrieval of relevant financial information from knowledge base
- Generation of personalized financial recommendations

Limitations:

- The system cannot execute financial transactions
- Recommendations are advisory in nature and not legally binding
- Performance depends on the quality of data sources and embedding models
- Limited to text-based interactions without visual or voice interfaces
- May not account for all regional financial regulations

1.5 Methodologies of Problem Solving

The project employs several methodologies to solve the problem of accessible financial advice:

1. **Modular Architecture:** Separating the system into specialized advisory modules
2. **Retrieval-Augmented Generation:** Combining information retrieval with generative AI
3. **Vector Similarity Scoring:** Using cosine similarity to match queries with relevant documents
4. **Hybrid Retrieval:** Combining semantic search with keyword-based retrieval (BM25)
5. **Prompt Engineering:** Creating effective prompts for the language model
6. **Evaluation Frameworks:** Using semantic similarity, ROUGE-L, and BLEU scores to assess output quality

02. Literature Survey

The literature survey reveals several important trends and gaps in financial advisory systems:

1. **Khandare et al.** explored a financial advisory system using Generative AI and RAG models with Google Bard, highlighting limitations of traditional robo-advisors and proposing context-aware chatbots.
2. **Kim et al.** investigated RAG implementation in financial AI, developing a multi-phase retrieval strategy to enhance information extraction from financial documents.
3. **Belanche et al.** studied robo-advisors in FinTech, examining adoption determinants including personal attitudes, mass media, and subjective norms.
4. **Hassnian et al.** created a hybrid system with retrieval processes and generative AI for the FinTech domain, outperforming legacy NLP models.
5. **Karangara** addressed integration of AI technologies in financial services, analyzing GANs and VAEs to enhance decision-making processes and customer experiences.

6. **Iaroshev et al.** assessed RAG models in financial report question-answering, finding that hybrid RAG models outperform isolated retrieval or generation solutions.
7. **Karangara** (in another study) examined Generative AI's impact on the financial sector, highlighting applications in customer relationship management, fraud detection, and trading.
8. **Pelau et al.** examined how perceived human-likeness in AI systems affects customer willingness to interact with GenAI.
9. **Jimeno-Yepes et al.** investigated improved chunking approaches for financial reports in RAG systems, finding that structural component-based chunking enhances retrieval accuracy.

Research Gaps Identified:

- Limited practical implementations of RAG in financial advisory contexts
- Absence of ethical and transparency frameworks
- Limitations in real-time data integration
- Lack of iterative development and feedback mechanisms
- Insufficient regulatory compliance considerations
- Need for hybrid AI and traditional advisory approaches

03. System Design

19 3.1 System Architecture

The financial advisory system uses a modular architecture consisting of several key components:

1. **Data Retrieval Module:** Collects financial data from various sources including web scraping and financial APIs (Alpha Vantage, Yahoo Finance, Bloomberg).
2. **Data Storage and Processing Layer:** Stores financial data in a structured manner using Google Cloud Storage and NoSQL databases. Performs data cleaning, preprocessing, and indexing.
3. **Knowledge Base:** Domain knowledge is categorized into individual PDF files according to financial categories (investment, insurance, loan services).
4. **Query Processing Module:** Processes user queries through Natural Language Processing to identify context, sentiment, and finance vocabulary.
5. **RAG Module:** Generates query embeddings and retrieves relevant contextual information from indexed sources.
6. **Generative AI Component:** Uses Google's Gemini-Pro model to generate coherent, personalized finance recommendations based on retrieved information.
7. **User Interface:** Provides a conversational interface for users to interact with the system.
8. **Advisor Engine Layer:** Contains three domain-specific modules (Loan Advisor, Investment Advisor, and Insurance Advisor) that process structured user input and provide customized recommendations.

23 The system uses a hybrid retrieval approach that combines vector similarity and BM25 keyword-based scoring to ensure both semantic and lexical relevance of retrieved information.

04. Project Implementation

4.1 Overview of Project Modules

The project consists of the following main modules:

1. General Finance Chatbot: Answers general finance queries using RAG approaches to fetch data from a pre-curated knowledge base.
2. Loan Advisory Bot: Accumulates user-specific data like credit score, purpose of loan, monthly income, and risk appetite to provide individualized loan suggestions.
3. Insurance Advisory Bot: Accumulates user information such as age, dependents, employment status, and medical conditions to suggest best insurance products.
4. Investment Advisory Bot: Examines investment horizon, financial goals, income, and risk tolerance to provide suitable investment advice.
5. Data Integration Module: Retrieves and processes financial data from different sources to make sure recommendations are up-to-date.
6. Vector Database: Keeps vectorized document chunks for similarity search efficiency in the retrieval phase.

4.2 Tools and Technologies Used

The project utilizes multiple tools and technologies:
1. Google Gemini Pro: Large language model for creating natural language output
2. Google's embedding-001 model: For embedding text and representing as vectors
3. Vector Databases: Chroma, Pinecone, or FAISS for similarity search with efficiency
4. Financial APIs: Alpha Vantage, Yahoo Finance, Bloomberg for financial data in real-time
5. Google Cloud Storage: To store structured financial data
6. NoSQL databases: For storing and retrieving data in flexible ways
7. Web Scraping Tools: For scraping financial data from trusted sources
8. Natural Language Processing (NLP): To process user input and derive relevant information4.3 Algorithm Details

4.3.1 Query Embedding and Document Embedding

This algorithm converts both user queries and document chunks into high-dimensional vectors using text embedding models:

```
Let Q be the user query
Let Di be a chunk of document text
Let fembed(·) be the embedding function

Then:
q⃗ = fembed(Q), d⃗i = fembed(Di)

Each document chunk Di is stored in the Vector Database with its corresponding
vector d⃗i
```

4.3.2 Vector Similarity (Cosine Similarity)

14 To find the most relevant documents for a query, cosine similarity measures the semantic closeness between the query vector and each document vector:

$$\text{sim}(\vec{q}, \vec{d}i) = (\vec{q} \cdot \vec{d}i) / (||\vec{q}|| \cdot ||\vec{d}i||)$$

Documents with the highest similarity scores are retrieved as relevant context for the query.

4.3.3 BM25 Keyword-Based Scoring

The system also performs keyword-based retrieval using the BM25 ranking function:

$$\text{BM25}(q, d) = \Sigma[i=1 \text{ to } n]\ \text{IDF}(qi) \cdot ((f(qi, d) \cdot (k1 + 1)) / (f(qi, d) + k1 \cdot (1 - b + b \cdot (|d|/\text{avgdl}))))$$

Where:
- f(qi, d) is the term frequency of term qi in document d
- |d| is the length of the document
- avgdl is the average document length in the corpus
- k1 and b are hyperparameters (commonly k1 = 1.2, b = 0.75)
- IDF(qi) is the inverse document frequency of term qi

4.3.4 Ensemble Hybrid Retrieval

The final set of relevant documents is derived using a weighted ensemble of vector similarity and BM25 scores:

$$\text{Score}(di) = \alpha \cdot \text{sim}(\vec{q}, \vec{d}i) + (1 - \alpha) \cdot \text{BM25}(q, di)$$

Where:
- $\alpha \in [0, 1]$ is a weight parameter controlling the balance between semantic and keyword search

05. Results

5.1 Outcomes

The financial advisory chatbot was evaluated using three widely accepted NLP metrics:

1. **Semantic Similarity:** Measures conceptual alignment between generated and reference responses.
2. **ROUGE-L:** Measures surface-level lexical overlap.
3. **BLEU:** Measures n-gram overlap between machine-generated text and human-written references.

The system was tested with five sample financial queries:

1. What is the best way to start investing with a small amount of money?
2. How should I allocate my 401(k) investments?
3. What's the difference between a traditional IRA and a Roth IRA?
4. How can I improve my credit score quickly?
5. Should I pay off my student loans early or invest that money?

Performance metrics summary:

- Average Semantic Similarity: 0.83

- Average ROUGE-L Score: 0.46
- Average BLEU Score: 0.24

These results indicate that the chatbot effectively understands user intent and provides contextually accurate responses, even if the exact wording differs from reference answers.

The system was also compared against prominent local LLMs including Mistral, LLaMA 2, and Gemma. Our RAG+Gemini model achieved superior scores across all metrics, with faster response times and better contextual relevance.

5.2 Screen Shots

[Note: This section would contain actual screenshots of the system interface and interactions, which are not available in the provided research paper.]

10

5.3 Evaluation Metrics and Graphical Analysis

The model's performance was evaluated using three key Natural Language Processing (NLP) metrics:

6

- **Semantic Similarity:** Assesses how closely the generated response matches the intended meaning of the reference answer. The model achieved an average Semantic Similarity score of **0.83**, with individual query scores ranging from **0.74 to 0.98**. These results indicate a strong ability to understand and preserve user intent.

2

- **ROUGE-L Score: Represents the longest common subsequence overlap between the output and reference response, indicating lexical similarity. ROUGE-L average score across the tested queries was 0.46, varying between 0.10 to 0.70. It indicates moderate surface overlap with strong semantic coverage.**

6

- **BLEU Score: Measures n-gram overlap between output and reference. The average BLEU score was 0.24, which reflects moderate similarity of lexicons but emphasizes the model's ability to produce varied but semantically correct language.**

The performances of these measures were illustrated graphically in Figures 4 to 8 in the form of bar charts, with comparative performance between various queries. The model achieved high semantic comprehension, moderate lexical similarity, and stable output quality, favoring meaning and user interaction over literal word matching.

16

06. Conclusion

6.1 Conclusions

The integration of Generative AI and Retrieval-Augmented Generation (RAG) in financial advisory systems is a major breakthrough in democratizing access to customized investment advice. The system developed surpasses traditional static robo-advisories by offering dynamic recommendations that adapt to specific user requirements and evolving market conditions. Major outcomes of the project are:

1. Successful combination of explainability and personalization, enabling users to see the rationale behind recommendations.
2. Application of behavioral finance concepts to mitigate cognitive biases that usually impair rational decision-making.
3. Multimodal data acquisition from multiple sources, augmenting recommendations with contextual data.
4. Design of a hybrid ethical AI system incorporating regulatory checks into the recommendation engine.
5. Design of a system that does not only act as a transactional engine but as a learning platform that enhances financial education through engaging dialogues.

The performance test attests that the RAG+Gemini method produces semantically correct and contextually appropriate financial guidance, beating the baseline language models on a range of metrics.

6.2 Future Work

A number of exciting areas for research and development in the future have been identified:

1. Quantum-Informed Adaptive Learning Models: Investigating quantum machine learning models to improve predictive power in volatile markets and to maximize portfolio investments.
2. Neuro-Symbolic AI for Dynamic Risk Profiling: Fusing neural networks with symbolic logic to assess risk dynamically, possibly with biometric signals.
3. Decentralized Autonomous Financial Agents (DAFAs): Creating blockchain-driven agents that are capable of adapting strategies autonomously through smart contracts.
4. Cross-Modal Sentiment Synthesis: Leveraging vocal tone, facial expression, and video-based analysis to identify early warning indicators of corporate instability or innovation.
5. Ethical AI Mirrors for Cognitive Bias Reduction: Designing systems that reveal to users the long-term impact of their decisions under different biases to promote better decision-making.
6. Self-Sovereign Identity (SSI) Integration: Rolling out systems that provide the user with complete control over their financial information, improving privacy and regulatory alignment.

These directions forward can make AI-based financial advisory systems into autonomous, morally principled collaborators that enhance investment results and overall financial health.

6.3 Applications

The implemented financial advisory system has the following real-world applications:

1. Personal Financial Planning: Assisting individuals in making well-informed decisions regarding savings, investments, and retirement planning.
2. Loan Decision Support: Helping users choose the right loan products according to their financial condition and requirements.
3. Insurance Selection: Advising users on ideal insurance coverage based on their individual conditions.
4. Investment Portfolio Management: Suggesting asset allocation and investment strategies according to risk tolerance and financial objectives.
5. Financial Education: Enhanced financial literacy where complex financial processes are simplified to easy-to-use language.
6. Retirement Planning: Assisting users in preparing for retirement with suitable savings and investment planning recommendations.
7. Debt Management: Offering guidance to manage and curb debt efficiently.
8. Financial Crisis Management: Providing assistance at times of economic downturn or domestic financial crisis.

The ease of access and capability of personalization of the system make it exceptionally useful for segments of people usually deprived of the benefits of expert financial guidance.

Appendix A: Problem statement feasibility assessment

Complexity Analysis of Financial Advisory System

The financial advisory system developed in this project can be analyzed for computational complexity to determine its feasibility. The core elements of the system include:

1. **Document Retrieval Problem**: Finding the most relevant documents for a user query
2. **Recommendation Generation Problem**: Generating optimal financial recommendations based on retrieved information

Document Retrieval Complexity

The document retrieval component of the system can be modeled as a nearest neighbor search problem in a high-dimensional vector space:

Given:

- A set of document vectors $D = \{d_1, d_2, ..., d_n\}$ in $\mathbb{R}^k$ (k-dimensional space)
- A query vector $q \in \mathbb{R}^k$
- A similarity function sim(q, d)

Find: The subset $S \subseteq D$ such that $\forall d \in S$, $\text{sim}(q, d) \geq \tau$ (where τ is a threshold)

This is an instance of the **Approximate Nearest Neighbor (ANN) problem**, which has the following complexity characteristics:

- **Brute force approach**: $O(n \cdot k)$ where n is the number of documents and k is the dimension of vectors
- **Using optimized data structures** (like KD-trees or Locality-Sensitive Hashing): $O(\log n)$ query time with $O(n \cdot k)$ preprocessing

The ANN problem is in class **P** (polynomial time), though the high dimensionality of document embeddings (typically 768 or 1024 dimensions) can make it computationally expensive in practice.

Recommendation Optimization Problem

The recommendation generation can be formulated as a constrained optimization problem:

Given:

- User profile U with constraints $C = \{c_1, c_2, ..., c_i\}$ (e.g., risk tolerance, time horizon)
- Set of possible financial instruments $F = \{f_1, f_2, ..., f_j\}$
- Utility function u(F, U) measuring the value of recommendation F for user U

Find: $F^* \subseteq F$ such that $u(F^*, U)$ is maximized subject to constraints C

This is an instance of a **Constrained Optimization Problem**, which can be reduced to the **Knapsack Problem** in certain formulations. The Knapsack Problem is NP-Hard, meaning no known polynomial-time algorithm exists to solve it optimally in all cases.

However, for practical applications in the financial advisory system:

1. The number of financial instruments considered is typically limited (< 100)
2. Greedy algorithms and heuristics provide near-optimal solutions
3. For many financial recommendations, we need good solutions rather than provably optimal ones

Satisfiability Analysis

The user query satisfaction can be modeled as a **Boolean Satisfiability Problem (SAT)**:

Given:

- A set of financial objectives $O = \{o_1, o_2, ..., o_m\}$ (e.g., high return, low risk)
- A set of financial recommendations $R = \{r_1, r_2, ..., r_n\}$
- A satisfaction function s(r, o) that returns True if recommendation r satisfies objective o

Find: A recommendation r^* such that $\forall o \in O, s(r^*, o) = \text{True}$

This is a SAT problem, which is NP-Complete. However, in our implementation:

1. The number of objectives is typically small (< 10)
2. Not all objectives need to be satisfied simultaneously (partial satisfaction is acceptable)
3. Modern SAT solvers can efficiently handle problems of this size

Feasibility Conclusion

Based on the complexity analysis:

1. The document retrieval component operates in polynomial time (P) and is feasible with proper indexing structures.
2. The recommendation generation component is NP-Hard in its general form, but practical constraints make it tractable:
   - Limited option space
   - Acceptable approximation algorithms
   - Bounded user constraints
3. The RAG architecture effectively decomposes the problem into manageable sub-problems:
   - Information retrieval (P complexity)
   - Language generation (linear in output length)

Therefore, the overall system is computationally feasible for real-world financial advisory applications, with response times that remain practical for interactive use. The theoretical complexity challenges are mitigated through algorithmic optimizations, efficient data structures, and appropriate problem formulation.

Appendix C: Plagiarism Report of project report

[This section would contain the plagiarism check results for the project report, which are not available in the provided material.]

References

[1] H. Zhu et al., "Implementing artificial intelligence empowered financial advisory services: A literature review and critical research agenda," Journal of Business Research, vol. 174, 2024, Art. no. 114494.

[2] S. Kim, H. Song et al., "Optimizing Retrieval Strategies for Financial Question Answering Documents in Retrieval-Augmented Generation Systems," Advances in Financial AI Workshop at the International Conference on Learning Representations (ICLR), 2025.

[3] D. Belanche et al., "Artificial Intelligence in FinTech: understanding robo-advisors adoption among customers," Industrial Management Data Systems, vol. 120, no. 9, pp. 1657–1674, 2020.

[4] H. Ali and A. F. Aysan, "What will ChatGPT Revolutionize in Financial Industry?," Social Science Research Network (SSRN), Oct. 2023.

[5] I. Iaroshev et al., "Evaluating Retrieval-Augmented Generation Models for Financial Report Question and Answering," Applied Sciences, vol. 14, no. 20, p. 9318, 2024.

[6] R. Akkiraju et al., "FACTS About Building Retrieval Augmented Generation-based Chatbots," arXiv preprint arXiv:2407.07858, 2024.

[7] R. Karangara et al., "Fintech's Generative AI Revolution: How AI is shaping the Future of Banking and Financial Services," International Research Journal of Modernization in Engineering Technology and Science, vol. 5, no. 12, pp. 1812–1813, Dec. 2023.

[8] Q. Yang et al., "Enhancing Financial Advisory Services with GenAI: Consumer Perceptions and Attitudes Through Service-Dominant Logic and Artificial Intelligence Device Use Acceptance Perspectives," Journal of Risk and Financial Management, vol. 17, no. 10, p. 470, 2024.

[9] A. Jimeno Yepes et al., "Financial Report Chunking for Effective Retrieval Augmented Generation," arXiv preprint arXiv:2402.05131v3, 2024.

[10] Xiao, M., Jiang, Z., Qian, L., et al. (2025). Enhancing Financial Time-Series Forecasting with Retrieval-Augmented Large Language Models.

[11] Velusamy et al., "AI-Powered Chatbots in Financial Services". Journal of Universal Science and Technology, 5(1), 15–22.

[12] Huang, A. et. al, (2024). "How to Build a Retrieval-Augmented Generation Chatbot". Anaconda Blog.

[13] Auquan et. al (2023). The Advantages of RAG AI (Retrieval Augmented Generation) Over Generative AI for Financial Services. Auquan Insights.

[14] Geniusee et. al (2024). RAG as a Service: Retrieval-Augmented Generation Solutions for FinTech. Geniusee Blog.

[15] HatchWorks et. al (2024). RAG in Financial Services: Use-Cases, Impact, Solutions. HatchWorks Blog.

[16] Nystrom, E. (2025). LLMs and RAG for Financial Data in the FinTech Domain. Umeå University.

[17] Lee, J. et. al (2024). Multi-Reranker: Maximizing Performance of Retrieval-Augmented Generation in the FinanceRAG Challenge.

[18] El-Shihy et. al (2024). The Influence of AI Chatbots in Fintech Services on Customer Loyalty within the Banking Industry. Finance, Business Accounting, 3(1), 16–28.

[19] Lakkaraju et. al (2023). "Can LLMs be Good Financial Advisors?: An Initial Study in Personal Decision Making for Optimized Outcomes". arXiv preprint arXiv:2307.07422.

[20] Sarmah B. et. al, (2024). "HybridRAG: Integrating Knowledge Graphs and Vector Retrieval Augmented Generation for Efficient Information Extraction. arXiv preprint arXiv:2408.04948.