FinChat

CMPE 258 Final Project, Milestone 4

Name of Leader: Pawan Aditya Man
Other team members: Amarender Reddy Jakka, Shreekar Kolanu

One sentence summary

Developing an AI-powered personal financial advisor by fine-tuning base model and RAG-integrated LLMs, evaluated with RAGAS for optimized, real-time financial advice.

Abstract

The FinChat project introduces a cutting-edge approach to building a personal financial advisor by combining a fine-tuned Large Language Model (LLM) with a Retrieval-Augmented Generation (RAG) model. Unlike proprietary models like BloombergGPT, which are costly and restricted, FinChat leverages publicly available data sources such as finance-related books and Wikipedia to create an accessible, legally compliant solution for financial guidance.

By integrating static, fine-tuned LLMs with dynamically adaptive RAG models, FinChat delivers personalized, real-time financial advice that adjusts to changing market conditions. This dual-model architecture bridges the gap between interpretability and contextual relevance, addressing the limitations of traditional static advisory systems. Inspired by the open-source principles of FinGPT, FinChat promotes transparency, collaboration, and democratized access to advanced financial tools.

The market for AI-driven financial advisory is growing rapidly as individuals and small businesses increasingly seek affordable, personalized insights. FinChat fills this gap by providing a cost-effective, adaptive solution that aligns with the rising complexity of financial decision-making and the digital transformation of the financial sector.

Introduction

Financial chatbots are transforming the way individuals and organizations manage their finances by offering personalized, accessible, and efficient advisory services. These AI-powered tools leverage advancements in natural language processing (NLP) and large language models (LLMs) to provide tailored insights into budgeting, investments, and financial planning. Unlike

traditional financial advisors or static robo-advisors, financial chatbots can dynamically integrate real-time data to address rapidly changing market conditions. However, proprietary models, such as BloombergGPT, rely heavily on restricted datasets and incur significant costs, making them inaccessible to a broader audience [1]. Open-source frameworks like FinGPT have demonstrated how financial AI models can democratize access to reliable advisory tools, encouraging transparency and innovation in the financial domain [1], [2].

The FinChat project builds on this foundation by proposing an open-source financial chatbot that delivers personalized financial advice using a dual-model framework. This approach combines fine-tuned Large Language Models (LLMs) for domain-specific expertise with a Retrieval-Augmented Generation (RAG) system to integrate real-time financial news and market updates dynamically. RAG frameworks have proven advantageous for long-form financial Q&A by integrating up-to-date data from external sources, enhancing response precision for complex financial inquiries [3].

By maintaining an updated vector database, FinChat ensures its recommendations remain timely and contextually relevant, addressing the limitations of static advisory systems [3].

A major challenge encountered during the project was data curation, as the chatbot required datasets to be meticulously modified and structured to ensure relevance and personalization. This step was crucial for developing a signature response style that reflects user-specific needs. To achieve this, multiple rounds of fine-tuning were necessary, which demanded substantial computational resources and incurred significant costs due to infrastructure requirements. Despite these challenges, FinChat highlights the potential of open-source financial AI to bridge the gap between advanced technology and accessible financial guidance for individuals and small businesses [1], [4].

By integrating fine-tuned LLMs with RAG and addressing key challenges such as data relevance and computational efficiency, FinChat aims to set a new standard for AI-driven financial advisory tools, fostering inclusivity and innovation in the financial sector [2].

Materials and Methods

Datasets

Dataset name	Contents with examples	Pros	Cons
Personal Finance QA Dataset	2000 Question and answers with format **Input**: Questions related to personal finance. **Output**: Corresponding detailed answers offering insights and	The dataset encompasses a wide range of personal finance topics including budgeting, investments, and credit management, serving as a valuable educational resource	The dataset's 2000 QA pairs offer a strong base, but its limited diversity across financial scenarios and demographics may affect its generalizability

	guidance.	for enhancing financial literacy.	across varied user groups.
FinTextQA	FinTextQA contains 1,262 QA pairs, with 1,022 pairs from finance textbooks, accounting for 80.98% of the dataset, and 240 pairs from policies and regulations, accounting for 19.02% of the dataset.	The FinTextQA dataset comprises expert-authored Q&A pairs from renowned finance textbooks and authoritative sources like HKMA, EU, and FR, covering a range of question types from conceptual explanations to numerical calculations and comparative analyses.	The dataset is not available for open source because legal department at HSBC has not granted permission.
finanacial-advisor-100	The dataset has 100 QA pairs in format of **Question:** Personal finance-related query posted by a user. **Answer:** Detailed response generated by an AI or expert, providing advice or information based on the question. **Text:** Additional narrative or context related to the question, giving further insight into the user's financial situation or concerns	**Evaluating financial advising algorithms** to see how well they understand and respond to complex personal finance queries.	The dataset is very small compared to other dataset which can lead to bias and also the dataset might not cover all geographical or personal scenarios equally.

	A	B
1	input	output
2	Does U.S. tax code call for small business owners to count business purchases as personal income?	I am going to keep things very simple and explain the common-sense reason why the accountant is right: Also, my sister in law owns
3	What are 'business fundamentals'?	From http://financial-dictionary.thefreedictionary.com/Business+Fundamentals The facts that affect a company's underlying value.
4	Business Investment Loss from prior year	You need to give specific dates! In the United States, you have three years to file an amended tax return. https://www.irs.gov/uac/N
5	Would the purchase of a car for a business through the use of a business loan be considered a business	You don't say what country you live in. If it's the U.S., the IRS has very specific rules for business use of a car. See, for starters at least
6	Why is â€œcheque cashingâ€ a legitimate business?	In my experience (in the US), the main draw of check-cashing businesses (like "CheckN2Cash" is that they will hold your check for a c
7	Valuing a small business to invest in	It should be pretty obvious that without knowing what sort of assets the company owns, and what sort of net earnings are being gen
8	Borrow money to invest in a business venture with equity?	It's clearly a risk, but is it any different than investing in your own business? Yes, it is different. If you own a business, you determine
9	Why can't online transactions be completed outside of business hours?	Generally, unless you're doing a wire transfer, bank transactions are processed in batches overnight. So the credit card company wor
10	Full-time work + running small side business: Best business structure for taxes?	A tax return is a document you sign and file with the government to self-report your tax obligations. A tax refund is the payment you
11	New vending route business, not sure how to determine taxes	You're not paying taxes three times but you are paying three different taxes (or more). Sales tax is a business expense, just like costs
12	How to treat miles driven to the mechanic, gas station, etc when calculating business use of car?	Alright, IRS Publication 463: Travel, Entertainment, Gift, and Car Expenses Business and personal use. If you use your car for both bu
13	Can a business refuse to take credit cards?	Businesses are free to decide what payment methods they accept for their goods and services. Businesses sometimes advertise what
14	Get a loan with low interest rate on small business	I am going to assume your location is the US. From what I am seeing it is unlikely you will get a loan other than some government ba
15	Using cash back rewards from business credit card	A C-Corp is not a pass-through entity, any applicable taxes would be paid by the Corporation, which is a separate legal entity from yo
16	Moving a personal business to a LLC accounting in California	You can move money in and out of the business at will, just keep track of every transaction. Ideally you'd use an accounting software
17	US Double Taxation - Business Trips and the Foreign Tax Credit	If you're a US citizen, money earned while in the US is sourced to the US. So you can't apply FTC/FEIE to the amounts attributable to
18	Can I use my Roth IRA to start a business?	Read the Forbes article titled IRA Adventures. While it's not the detailed regulations you certainly need, the article gives some great

Personal Finance QA Dataset.

Models

Model name	Description with model diagrams/results	Classification accuracy (top 1)	Pros	Cons	Notes
FinGPT [1]	Open-source large language model tailored for financial applications with RAG capabilities for real-time data integration	Median: 80%, Max: 85%	High adaptability, open-source, access to updated financial data	Complex setup, reliance on high-quality financial datasets	Focuses on democratizing access to financial data and tools; ideal for research and innovation in open finance
FinTextQA [2]	RAG-based LFQA model, combining embedding, retrieval, reranking, and generation for comprehensive financial question answering	Median: 75%, Max: 82%	Effective for complex queries, integrates real-time data retrieval	Requires extensive processing, limited open datasets	Excels in long-form financial question answering, tackling domain-specific terminologies
BloombergGPT [4]	Proprietary financial LLM trained on Bloomberg's private dataset, optimized for finance-specific NLP tasks such	Median: 87%, Max: 92%	High accuracy, proprietary data sources ensure financial relevance	Not open-source, high cost of implementation	Designed specifically for Bloomberg's ecosystem, with exclusive access to financial information

Logistic Regression	Traditional statistical model commonly used for binary classification tasks in finance (e.g., credit risk analysis)	Median: 65%, Max: 72%	Lower compute power required, easily interpretable, straightforward to implement	Limited in handling complex, non-linear dependencies in finance	Suitable for basic financial classifications but lacks the depth of LLMs or deep learning models for NLP tasks

In summary, in terms of classification accuracy:

- The best-performing model is BloombergGPT with a maximum accuracy of 92%.
- The median performing model is FinGPT with a median accuracy of 80%.
- The worst-performing model is Logistic Regression with a maximum accuracy of 72%.

However, there are other performance metrics to consider. E.g., Computing power, Training Time, Memory Usage, User Satisfaction.

- The model with the least compute requirement is Logistic Regression.
- The model with the highest compute requirement is BloombergGPT.

Work Plan

Fine-Tuning the Model

1. **Dataset Preparation**:
   - A dataset of 2000 financial questions and answers was curated and formatted into JSONL as required by OpenAI's fine-tuning framework.
   - Each dataset entry consists of:
     - **System Message**: Instructions guiding the tone and structure of responses.
     - **User Message**: Financial queries curated from reliable sources.
     - **Assistant Response**: Targeted answers aligned with the system's goals.

```
1 {"messages": [{"role": "system", "content": "Marv is a factual
2 {"messages": [{"role": "system", "content": "Marv is a factual
3 {"messages": [{"role": "system", "content": "Marv is a factual

sarcastic."}, {"role": "user", "content": "What's the
sarcastic."}, {"role": "user", "content": "Who wrote '
sarcastic."}, {"role": "user", "content": "How far is

of France?"}, {"role": "assistant", "content": "Paris, as if
d Juliet'?"}, {"role": "assistant", "content": "Oh, just som
 from Earth?"}, {"role": "assistant", "content": "Around 384
```

```
instructions_string ="""FinChat is a AI chat assistant primarly focusing on personal finance advices for users,
addresses user queries on topics ranging from investment strategies, Budgeting, Saving, Retirement Planning,
Insurance and Risk Management,Credit and Debt Management. \
It responces to questions fom user and concludes with its signrature ' -FinChat'. \
It uses data which was used for fine tuning the model to answer and question for better responce\

Here are examples of FinChat responding to user questions.
```

2. **Data Splitting**:
   - The dataset is split into an 80:20 ratio for training and validation sets, ensuring the model is evaluated effectively during fine-tuning.
3. **Fine-Tuning Process**:
   - Fine-tuning was carried out using OpenAI's ChatGPT-3.5 Turbo API, which provides a supervised fine-tuning method to specialize the model for financial advisory tasks.
   - The resulting fine-tuned model is optimized for producing structured, concise, and personalized responses to financial queries.

```
test_question = "What are some effective budgeting techniques?"
```

```
Here's a variant suggestion assuming you're employed:
Try to compute and maintain a comprehensive income statement that includes amortizations for bigger expenses.
At the beginning of the month, when you compute your income, you should also add to it those amortizations to calculate an
effective income, from which you'd calculate your spending limits.
"Amortization" is a variation from what is known as "zero-based budgeting" where, in order to work out a budget, you sit down and
try to write out a comprehensive list of things you want to (and have to) spend for, and you can only start the month after you've
worked this out.
The problem I see with that, is that if you need to spend $X/mo for as long as you can see into the future (e.g. you have to drive
to work, etc.), then the derived $X/mo becomes the minimum you can ever manage with.
The added complication is that your actual income tends to see some slight variation over the year, if bonuses come and go, you
have vacation days, etc. So, with the effective income x(t) (functions added for the somewhat randomness from year to year), at the
beginning of time there is: x(0) - x(0) = 0, so you can't spend, anyway, unless you already have a buffer (and you probably do).
In recursion, with spending having dropped your effective wealth below x(t-1), you can now spend only up to x(t). -FinChat
```

4. **Performance Validation**:
   - Evaluation metrics, such as BLEU and ROUGE scores, were employed to measure the accuracy and quality of the model's responses to diverse financial scenarios.
5. **Challenges**:
   - Fine-tuning required significant computational resources, which posed challenges in GPU availability and infrastructure management.
   - OpenAI's fine-tuning services helped mitigate these issues by reducing dependency on local infrastructure.

Retrieval-Augmented Generation (RAG) Implementation
1. **Base Model**:
   - The base model for the RAG implementation is **Meta/LLaMA-3.1-405B-Instruct**, selected for its high adaptability to domain-specific tasks and real-time data retrieval capabilities.
2. **Document Embedding and Vector Database**:
   - Financial documents, such as PDFs and market reports, are divided into smaller chunks using **PyPDFLoader** and **RecursiveCharacterTextSplitter**.
   - Each chunk is embedded using NVIDIA embedding models and stored in a **FAISS vector database**, facilitating efficient similarity-based retrieval.

3. **HyDE Retriever Workflow**:
   - o **Initial Step**: The user's question is passed to the base model (**Meta/LLaMA-3.1-405B-Instruct**), which generates a hypothetical document or response summarizing the context of the query.
   - o **Information Retrieval**: Based on the generated hypothetical response, the system queries the vector database to retrieve contextually relevant information.
   - o **Final Response Generation**: The retrieved information, the user's original question, and a predefined instruction string are passed through LangChain to generate a final, comprehensive response.
4. **LangChain Integration**:
   - o The LangChain framework facilitates the interaction between the base model, vector database, and final response generator.
   - o Structured prompts guide the retriever and generator, ensuring that the context retrieved from the database is integrated seamlessly into the response generation process.
5. **Testing the RAG Model**:
   - o Sample queries are tested to validate the system's ability to retrieve and integrate relevant financial data dynamically.
   - o Example: A query on "Five shiny marketing tactics used by banks" involves extracting and summarizing relevant details from embedded documents, ensuring the response is accurate and contextual.

Results

We have used various metrics for our model like we have used the BLEU and ROUGE for the Fine Tuned model evaluation and then faithfulness, answer_correctness, context_precision, context_recall, answer_relevancy these for the evaluation for the Rag Imtegrated model.

BLEU:
Measures the overlap of n-grams between the predicted and ground truth texts. The sentence_bleu function computes BLEU for each prediction.
ROUGE:
ROUGE-1: Measures overlap of unigrams (single words).
ROUGE-2: Measures overlap of bigrams (two consecutive words).
ROUGE-L: Measures the longest common subsequence overlap.

Table 1: ChatGPT-3.5 Turbo Fine-Tuned Model Results

Metric	Score
BLEU Score	0.0184
ROUGE-1 Score	0.3239
ROUGE-2 Score	0.0519
ROUGE-L Score	0.1270

Faithfulness is the measure the factual consistancy of the generated answer against the given context.

Answer_correctness is the accuracy of the generated answer when compared to the ground truth.

Context_precision Evaluates the precision of the context used to generate an answer, ensuring relevant information is selected from the context.

Context_recall Measures if all the relevant information required to answer the question was retrieved.

Answer_relevancy Assesses how pertinent the answer is to the given question. It is validating that the response directly addresses the user's query.

Table 2: RAG Model Results (Example 1 and Example 2)

Metric	Example 1 Score	Example 2 Score
Faithfulness	1.0	1.0
Answer Correctness	0.4250	0.4250
Context Precision	0.9999	0.9999
Context Recall	1.0	1.0
Answer Relevancy	0.9462	0.9462

Sample Results

Results for ChatGPT 3.5 Turbo Fine tuned model:

```
test_question = "What are some effective budgeting techniques?"
```

```
Here's a variant suggestion assuming you're employed:
Try to compute and maintain a comprehensive income statement that includes amortizations for bigger expenses.
At the beginning of the month, when you compute your income, you should also add to it those amortizations to calculate an
effective income, from which you'd calculate your spending limits.
"Amortization" is a variation from what is known as "zero-based budgeting" where, in order to work out a budget, you sit down and
try to write out a comprehensive list of things you want to (and have to) spend for, and you can only start the month after you've
worked this out.
The problem I see with that, is that if you need to spend $X/mo for as long as you can see into the future (e.g. you have to drive
to work, etc.), then the derived $X/mo becomes the minimum you can ever manage with.
The added complication is that your actual income tends to see some slight variation over the year, if bonuses come and go, you
have vacation days, etc. So, with the effective income x(t) (functions added for the somewhat randomness from year to year), at the
beginning of time there is: x(0) - x(0) = 0, so you can't spend, anyway, unless you already have a buffer (and you probably do).
In recursion, with spending having dropped your effective wealth below x(t-1), you can now spend only up to x(t). -FinChat
```

Results for Rag Integrated Model:

Example 1:

Question: FIVE SHINY MARKETING TACTICS used by BANKS

Retrieved Context: [Document(metadata={'source': '/content/drive/MyDrive/data/i_will_teach_you_to_be _rich.pdf', 'page': 127}, page_content='FIVE SHINY MARKETING\nTACTICS BANKS USE TO TRICK\nYOU\n1. TEASER RATES ("6 percent for the first two months!"). Don't\nget sucked in by this trick—your first two months don't matter.\nYou want to pick a good bank that you can stick with for years—\none that offers overall great service, not a promo rate that will earn\nyou only $25 (or, more likely, $3). Banks that offer teaser rates are,\nby definition, to be avoided.\n2. REQUIRING MINIMUM BALANCES to get "free" services'), Document(metadata={'source': '/content/drive/MyDrive/data/i_will_teach_you_to_be _rich.pdf', 'page': 127}, page_content='like checking and bill paying. No, I'm not going to agree to a\nminimum amount. I'll just go somewhere else.\n3. UP-SELLS to expensive accounts ("Expedited customer\nservice! Wow!"). Most of these "value-added accounts" are\ndesigned to charge you for worthless services. I can't wait to have\nkids one day so my three-year-old can walk into a Wells Fargo,\nthrow his lollipop at the bank manager, and say "THIS ACCOUNT\nIS CLEARLY A RIP-OFF!" Good job, little Raj.'), Document(metadata={'source': '/content/drive/MyDrive/data/i_will_teach_you_to_be _rich.pdf', 'page': 126}, page_content='which are not by seeing how straightforward they are with their accounts\nand fees. Your bank shouldn't nickel-and-dime you through minimums\nand fees. It should have a website with clear descriptions of different\nservices, an easy setup process, and 24/7 customer service available by\nphone. Another thing: Ask them if they send you promotional material\nevery damn week. I don't want more junk mail! I don't need more cross-\nsells! I actually switched my car insurance because they would not stop'), Document(metadata={'source': '/content/drive/MyDrive/data/i_will_teach_you_to_be _rich.pdf', 'page': 127}, page_content='IS CLEARLY A RIP-OFF!" Good job, little Raj.\n4. HOLDING OUT by telling you that the no-fee, no-minimum\naccounts aren't available anymore. They are. Banks will resist\ngiving you a no-fee, no-minimum account at first, but if you're\nfirm, they'll give you the account you want. If they don't, find\nanother bank. There are many, many choices, and it's a buyer's\nmarket.\n5. BUNDLING A CREDIT CARD with your bank account. If\nyou didn't walk in specifically wanting the bank credit card, don't\nget it.')]

Model Response: According to the context, the FIVE SHINY MARKETING TACTICS used by BANKS to trick customers are:

1. TEASER RATES (e.g., "6 percent for the first two months!")
2. REQUIRING MINIMUM BALANCES to get "free" services
3. UP-SELLS to expensive accounts (e.g., "Expedited customer service! Wow!")
4. HOLDING OUT by telling customers that no-fee, no-minimum accounts aren't available anymore
5. BUNDLING A CREDIT CARD with a bank account when the customer didn't specifically ask for it.

Exaample 2:

Question: does money makes us happy?

Retrieved Context: [Document(metadata={'source': '/content/drive/MyDrive/data/i_will_teach_you_to_be _rich.pdf', 'page': 192}, page_content='drinking instant\ncoffee and packing\nmy lunch.\n—JAMES CAVALLO, 27\nMy Rich Life is guilt-free spending. I no longer say I can't afford\nX, Y, or Z. I say I choose not to spend on it.\n—DONNA EADE, 36\nDOES MONEY MAKE US\nHAPPY?\nYes! I know, I know. You may have heard about a study that found\nmoney makes us happy up to $75,000, then it levels off. In reality,\nthe 2010 study by Deaton and Kahneman found that "emotional\nwell-being" peaks at $75,000. But if you take another measure,'), Document(metadata={'source': '/content/drive/MyDrive/data/i_will_teach_you_to_be _rich.pdf', 'page': 192}, page_content='"life satisfaction," you find no plateau—not at $75,000, or\n$500,000, or even $1 million.\nAs Dylan Matthews notes in an excellent Vox article, there is\nstrong data indicating that the more you earn, the more satisfied\nyou are with your life. "For developing and developed countries\nalike, being richer is correlated with higher life satisfaction."\nAnd if you want to know how to use money to live a happier\nlife? Whillans et al. told the New York Times that "People who'), Document(metadata={'source': '/content/drive/MyDrive/data/i_will_teach_you_to_be _rich.pdf', 'page': 35}, page_content='value different things. Paul loves eating out at Michelin-starred\nrestaurants where a meal might cost $500. Nicole loves traveling. And\nNick loves buying clothes. If you don't consciously choose what "rich"\nmeans, it's easy to end up mindlessly trying to keep up with your friends.\nI consider myself rich now that I can do these things:\n■ Make career decisions because I want to, not because of money\n■ Help my parents with their retirement, so they don't have to work if\nthey don't want to'), Document(metadata={'source': '/content/drive/MyDrive/data/i_will_teach_you_to_be _rich.pdf', 'page': 211}, page_content='For me, it's eating out, travel, and clothes. Depending on the time of\nyear—or how nice a cashmere sweater I found—those numbers can vary\nby thousands of dollars a month.')]

Model Response: According to the context, yes, money can make us happy. A 2010 study by Deaton and Kahneman found that "emotional well-being" peaks at $75,000, but "life satisfaction" continues to increase with higher earnings, with no plateau. Additionally, research by Whillans et al. suggests that people who use money to live a happier life tend to value different things and make conscious choices about what "rich" means to them.

Discussion and Conclusion

We compared the performance of the ChatGPT-3.5 Turbo Fine-Tuned Model and the RAG Model. The results from the fine-tuned model highlight that fine-tuning is effective for structuring responses and focusing on key parameters. However, the fine-tuning process yielded low scores, primarily due to limited computational resources that restricted multiple fine-tuning iterations with varying instructions. Additionally, the quality and structure of the dataset played a significant role in influencing the model's responses.
The RAG model, on the other hand, performed significantly better, especially in retrieving and integrating relevant context for dynamic and real-time financial advisory tasks. For future work, we aim to implement a robust data pipeline to ingest up-to-date data into the vector database. This enhancement would improve the RAG model's ability to generate responses that are both accurate and reflective of the latest information.

Market and Benefit

Market Demand for AI Financial Advisors:
- Increasing Adoption in Financial Services, Financial institutions, like Citigroup, are integrating AI tools to streamline operations and enhance productivity in advisory services. [Citigroup rol](#)
- Clients increasingly expect tailored financial advice, and AI empowers advisors to provide personalized investment solutions for navigating complex markets more effectively. [A Changing Fina](#)

Benefits of Open-Source Financial AI Models:
- Improved Financial Literacy: Offers accessible and personalized financial advice, empowering users to make informed decisions about budgeting, investing, and saving.
- Market Accessibility: Makes advanced financial tools accessible to small businesses and underserved populations, promoting inclusivity in financial services.
- Real-Time Insights: Provides up-to-date financial information, helping users adapt to dynamic market conditions and seize opportunities.
- Benchmarking New Architectures: Evaluates the effectiveness of fine-tuned LLMs and RAG models, contributing valuable insights to AI and financial technology research.
- Cost-Effective Advisory: Eliminates the need for costly human financial advisors, making expert guidance available to a wider audience.
- Market Accessibility: Makes advanced financial tools accessible to small businesses and underserved populations, promoting inclusivity in financial services.

Alternates/competitors

Alternate	Advantages of Our Approach	Notes
Proprietary Financial AI Models (e.g., BloombergGPT)	* Open-source and more accessible to individuals and small businesses. * Lower cost compared to proprietary solutions.	Requires curation of public datasets for compliance.
Static Financial Advisory Systems	* Real-time data retrieval via RAG ensures up-to-date and accurate responses. * Personalized advice tailored to user queries.	Limited in adapting to dynamic market trends.
General-purpose Chatbots (e.g., ChatGPT)	* Fine-tuned for financial-specific tasks. * Enhanced contextual relevance through vector database integration.	Relies on financial dataset curation.

Appendix

This is the link to our private Github repo: https://github.com/pawanaditya85/FinChat

References
[1] Y. Nie et al., "A Survey of Large Language Models for Financial Applications: Progress, Prospects and Challenges," arXiv preprint, 2024.
https://arxiv.org/html/2406.11903v1

[2] J. Chen et al., "FinTextQA: A Dataset for Long-form Financial Question Answering," arXiv preprint, 2024.
arXiv:2306.06031v1.

[3] C. Fieberg, L. Hornuf, and D. J. Streich, "Using Large Language Models for Financial Advice," arXiv preprint, 2024.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4850039

[4] Wu et al., "BloombergGPT: A Large Language Model for Finance," 2023.
https://arxiv.org/abs/2303.17564

LangChain Documentation LangChain Framework for Context-Aware Response Generation.
LangChain Official Docs, 2024. [Online]. Available: https://langchain.readthedocs.io/

RAGAS Library "Retrieval-Augmented Generation Evaluation Toolkit."
GitHub Repository, 2024. [Online]. Available: https://github.com/explodinggradients/ragas

https://www.reuters.com/technology/artificial-intelligence/citigroup-rolls-out-artificial-intelligence-tools-employees-eight-countries-2024-12-04/?utm_source=chatgpt.com

https://www.forbes.com/sites/forbes-shook/2024/05/30/a-changing-financial-world-artificial-intelligence/?utm_source=chatgpt.com