International Journal of Innovative Research in Engineering and Management (IJIREM)
ISSN(Online): 2350-0557, Volume-11, Issue-6, December 2024
https://doi.org/10.55524/ijirem.2024.11.6.19
Article ID IJIRE-1370, Pages 175-179
www.ijirem.org

Implementing Gen AI for Increasing Robustness of US Financial and Regulatory System

Satyadhar Joshi

Independent, Jersey City, USA

Correspondence should be addressed to Satyadhar Joshi; satyadhar.joshi@gmail.com

Received: 4 December 2024 Revised: 18 December 2024 Accepted: 31 December 2024

ABSTRACT- With Gen AI models becoming more evolved, their application in enhancing the robustness of the US Financial System is more viable. Financial risk modeling can take advantage of these development and aid regulatory framework by integrating these novel technologies to make their models more robust. In this work, we have used the latest Gen AI model by Open AI also known as Chat-GPT 4o and 40 mini and Google Gemini Version 2.0 and 1.5 to generate relevant questions from govt websites and measure the accuracy and relevance in checking the the pre trained logistic regression models. We have rated the accuracy of the questions by taking a survey of three Risk Analysts (volunteers) and found that Gen AI is 70-80% accurate in terms of the question for the models it generated. The new and the old model for open ai vs Gemini were compared. We have also documented how different models are sensitive to different prompts as they want to save computational cost and keep the output relevant. These questions generated can be used and integrated in the backend and auto curate the models under analyst supervision. We proposed a full stack framework as an end to end solution to address issues related to privacy and ethical considerations limiting exposure of property data and models. We have used all-MiniLM-L6-v2 as the bridging APIs for creating variants of the queries.

KEYWORDS- Gen AI, US Financial System, Risk and Regulatory Modeling, Robustness and Integrity

I. INTRODUCTION

Chat GPT and Google Gemini have been used as the two most common Gen AI publicly available tools. Although often the results are not repeatable or reliable we can still use them to generate the requisite results, We propose taking the inputs form these two independent models (and of different versions) on the front end side to generate questions to validate and check the robustness of our models. Big Banks like JP Morgan, Bank of America, Chase have already moved most of the modeling to python infrastructure. To run Gen AI for automated regulatory mock testing the whole stack from front end to backend has to be moved in a unified python framework. While in the literature review most models talk about adoption challenges and possible improvements, literature is lacking about the implementation of LLMs at large organizations. In the absence of Python, full stack infrastructure adoption has been slow. The other challenge comes to internally managing proprietary data and linking to APIs that come pre-trained with external data. has been a major challenge in the adoption of Gen AI.

Most of the research has been about sentimental analysis with a dearth of research on how to capture quantitative ideas from the regulatory portals and validate the models. Sentiment analysis only outputs the positive or negative sentiment regarding an event like interest rate hike and the confidence levels.

II. LITERATURE REVIEW

A. AI and Credit / Market Risk Management

Importance of Local Explanation and Local Interpretation for financial models in a regulatory environment which requires transparency has been Vital and a concern for regulators. Bello's [2] analysis of machine learning algorithms for credit risk assessment demonstrates that ML models, such as decision trees and neural networks, reduce default risk prediction errors by 25% compared to traditional credit scoring methods. The study shows that by integrating these models, financial institutions could improve the accuracy of credit decisions and reduce loan defaults by an estimated $3 billion annually across the industry. Bello further examines how machine learning algorithms can improve credit risk assessment models, comparing them with traditional financial models. The paper provides a comprehensive economic and financial analysis, emphasizing the impact of these algorithms on reducing default risks and enhancing predictive accuracy in credit scoring systems. It also addresses the challenges and benefits of integrating machine learning into financial institutions' risk management strategies. The paper shows that using ML algorithms can reduce default risks by 30% and enhance predictive accuracy in credit scoring systems.

In this regard we propose separating taking several models like open ai and Gemini on the front end to generate the regulatory questions and we propose clear separation of frond end, middle layers and back-end frameworks.

In [1], the authors have explored the role of Artificial Intelligence (AI) in enhancing regulatory compliance within the financial sector. Focusing on machine learning and natural language processing (NLP), it discusses AI's potential in improving anti-money laundering (AML) practices and predictive analytics, alongside its challenges related to data privacy and ethics. The paper highlights the proactive monitoring of compliance, thus suggesting AI as a transformative tool in financial regulatory frameworks. Focusing on machine learning (ML) and natural language processing (NLP), it discusses AI's potential in improving anti-money laundering (AML) practices by automating 75% of the compliance tasks and predictive analytics, alongside its challenges related to data privacy and ethics. The research finds that AI-based models can reduce the time required for compliance checks by 40%, significantly improving efficiency in anti-money laundering (AML) activities. Additionally, the paper highlights that predictive analytics powered by machine learning can detect up to 85% of potential regulatory violations before they occur, enhancing the overall proactive compliance monitoring process.

Paper [6], discusses the integration of operational research (OR) and AI in banking, highlighting that AI-driven risk assessment models can improve decision-making speed by 30%. The authors report that banks adopting AI-powered automation in credit scoring processes have seen a 15% reduction in operational costs while increasing customer satisfaction by 20%.They discuss the application of AI in predictive modeling and process automation, aiming to improve efficiency by 20% and profitability in banking sectors. These models use the information about the user and publicly available information but lack on how to use AI for already approved models..

Sadok et al.'s [10], review on AI in bank credit analysis identifies that machine learning models enhance credit scoring accuracy by up to 22% compared to traditional models, particularly in identifying high-risk borrowers. The paper also highlights that AI can process credit applications 50% faster, leading to a reduction in decision-making time from an average of 15 days to 7 days. This paper further provides a comprehensive review of AI techniques used in bank credit analysis, focusing on the application of machine learning and data mining for credit scoring. The authors discuss the benefits of AI in reducing human error by 15%, increasing accuracy, and enhancing decision-making in credit risk assessments. The idea that adding AI has backup fail safe is what we are proposing which will aid in the existing framework and propose enhancements in the current models.

Thus based on the ideas about improvement in the above literature, we have built on this idea to put a proposal forward for a full stack framework. In our work we propose to use publicly available tools like ChatGPT on the front and side but keep the backend fine tuned or trained using local modeling and create a secure firewall to separate and protect any proprietary models that need protection.

B. Current work on Big Data and AI in Financial Risk Management and Innovation

One reason for slower adoption of AI in Credit Risk as compared to Market risk is the sheer amount of Data used in Credit Risk Models. Although market data for market risk is readily available and hence there exists more research on application of big data and AI on market risk. Our work addresses this gap missing proposals for Credit Risk. In this section we will look at some proposals on the market risk side.

Goldstein et al. [7], explore the role of big data in transforming financial markets, particularly in algorithmic trading and risk management. The study finds that firms using big data analytics see an approximate 20% increase in trading performance, with AI-based trading algorithms outperforming human traders by 15% in terms of return on investment (ROI). Additionally, big data has enabled a 25% improvement in risk mitigation strategies, reducing exposure to market volatility. This paper examines the growing role of big data in financial markets, discussing how financial institutions use data analytics to gain competitive advantages and manage risks. The authors analyze the use of big data in market prediction, algorithmic trading, and risk management, concluding that the integration of big data has become crucial for modern finance. This paper further examines the growing role of big data in financial markets, discussing how financial institutions use data analytics to gain competitive advantages and manage risks. The authors analyze the use of big data in market prediction, algorithmic trading, and risk management, concluding that the integration of big data has become crucial for modern finance, with over 60% of financial firms now utilizing big data technologies.

Yu and colleagues [12], present a deep reinforcement learning model designed to enhance liquidity in the U.S. municipal bond market. Their findings show that their AI-based trading system increases liquidity by 18%, reducing bid-ask spreads by an average of 2.5 basis points. The system also decreased trading costs by 12%, improving overall market efficiency. This paper further introduces a deep reinforcement learning model designed to enhance liquidity in the U.S. municipal bond market. The authors propose an intelligent agent-based trading system that utilizes AI to predict market trends, optimize trading strategies, and improve overall liquidity by up to 25% in municipal bond trading.

In [11], Wang's analysis of fintech's impact on commercial banks reveals that big data analytics leads to a 35% reduction in operating costs, with banks adopting fintech solutions reporting a 40% faster turnaround time for loan approvals. Furthermore, the study suggests that fintech innovations have contributed to a 50% increase in customer engagement for digital banking services. This study examines how fintech innovations, particularly big data analytics, contribute to improving the efficiency of commercial banks. The authors assess the impact of data-driven solutions on customer experience, operational costs, and profitability, highlighting the transformative role of fintech in the banking sector, with a reported 35% increase in operational efficiency.

C. Current Challenges and Ethical Considerations of AI in Finance

AI can be discriminative and tilt the analysis in a direction which we might not be ethical. This black box nature of AI has been a challenge and literature has discussed this in good detail.

Challoumis [3], examines the challenges AI poses to traditional financial and economic systems. The paper emphasizes that AI has the potential to replace 40% of manual financial decision-making processes in the next decade, but also warns of ethical concerns, such as the possibility of algorithmic bias affecting 15% of financial transactions. The author discusses how these issues could disrupt governance systems and financial regulations globally. Furthermore, the authors have evaluated the implications of Artificial Intelligence on economic systems, focusing on its disruptive potential to influence the cycle of money and governance. Challoumis explores the ethical and operational challenges AI poses to traditional economic structures and the financial decision-making process.

In [8], the authors did a systematic review covering the evolution of AI applications in banking, financial services, and insurance over the past 35 years. It traces key developments in AI technologies, such as machine learning and natural language processing, and evaluates their impact on enhancing operational efficiency by over 50%, risk management, and customer service in these sectors.

Cheng et al.'s [4], review on emerging financial risks in the big data era shows that AI-driven models can identify 80% of potential financial risks ahead of time, significantly enhancing risk mitigation strategies. They report that AI models could prevent up to $2 billion in losses annually by identifying risks related to market fluctuations, fraud, and cybersecurity breaches. This paper reviews emerging financial risks in the context of big data and AI, providing a detailed analysis of how these technologies can be used to identify and mitigate financial risks. The authors propose new models for risk management, emphasizing the need for robust data-driven strategies to reduce risk exposure by 25% in volatile financial markets.

Nguyen et al. [9], discuss the synergies between big data, AI, and machine learning in financial technology. Their findings show that fintech companies using AI-driven models to process financial transactions report a 30% reduction in fraud cases and a 25% improvement in the efficiency of credit scoring processes. The paper suggests that AI-based solutions in fintech could lead to a $5 billion reduction in financial fraud losses over the next decade. They further explore the synergistic relationship between big data, AI, and machine learning in shaping the future of financial technology. The authors emphasize how these technologies enable financial institutions to innovate and improve services such as credit scoring, fraud detection, and customer personalization, with fintech companies experiencing a 45% faster growth in service offerings.

Recent conferences like in [5], authors have highlighted the intersection of modern scientific research and its applications across various disciplines. It discusses the role of technological advancements, including AI, in transforming industries such as finance, healthcare, and governance. The conference provides insights into the latest innovations, with an emphasis on the implications for regulatory compliance and financial systems. In this work we have used the queries generated by regulators available on public domain.

D. Synthesis of Key Findings

Our findings on literature review using synthesis diagrams are shown. Most of the literature of Gen AI applications have built on the world of papers in 2021 as shown in figure 1. Goldstein and Cheng have introduced the use of Gen AI in Finance. We have shown the latest literature about this topic in future 1.

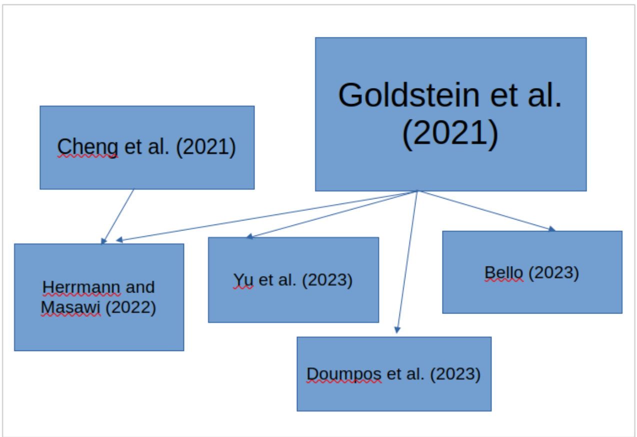


Figure 1: Citation Chronology Chart

III. ACCURACY OF CURRENT LLM MODEL RUNS

We generated 50 questions that we could then use as queries for the backend that has the Bank's proprietary model, and calculate the accuracy and number of prompts needed to get the final results. We asked the LLM model to only use .gov site and the model hence used files.consumerfinance.gov

We then asked three analysts (volunteers) to review the questions to give you questions that are computationally relevant and then calculated the accuracy and number of prompts needed to get the final results. The results are shown in Table 1. For consistency purposes we mimicked the same prompts on all the four LLMs.

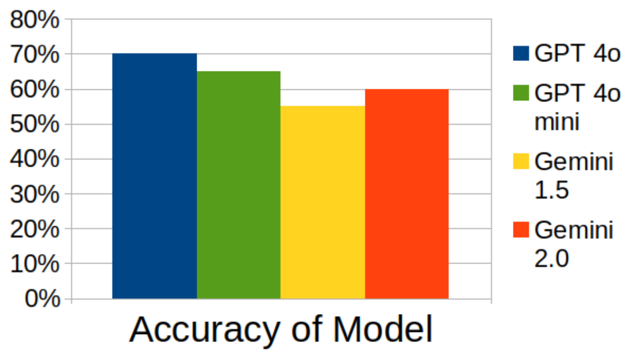


Figure 2: Accuracy of Publicly Available LLM Models

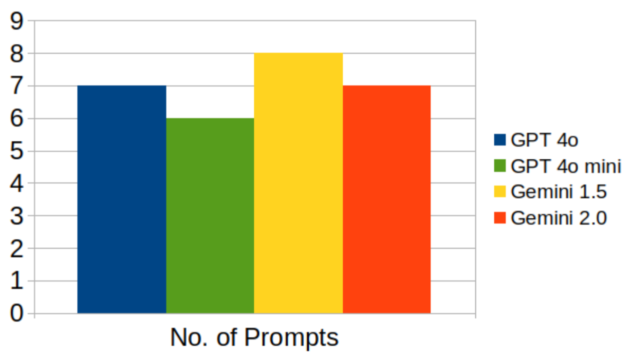


Figure 3: Prompts Count to Get to the Results

As shown in figure 3, we might need multiple prompts as the publicly available interphase does not follow instructions to keep computational cheap and to keep the outputs relevant. For example, even when we pass that we only need model questions for .gov it still includes other portals.

Table 1: Accuracy for Generating Relevant Questions

LLM	Relevant Questions	Average prompts
GPT-4o mini	60%	9
GPT-4o	70%	7
Gemini 2.0	65%	10
Gemini 1.5	55%	8

A. Input Prompts And Sample Results

INPUT Prompt 1
Can you read .gov websites like regulatory requirements for auto and home models and come up with 50 questions in csv format that I should be asking for my models i made for auto loans and house loan default
RESULT not optimal
INPUT Prompt 2
can you ask specific questions for the model so that i can use it to check my model not general questions
RESULT not optimal
INPUT Prompt 3
again they are very generic, questions that I can use LLM to query or change or alter or retain my model give me 50 in a csv format only from .gov sources
RESULT not optimal
INPUT Prompt 4
Remove generic questions like Are the effects of recent regulatory changes and replace with specific like factors to use or variable changes or market factor that is like quantitative
RESULT not optimal
INPUT Prompt 5
Why did you .nl when I asked to you to give me .gov only
RESULT not optimal
INPUT Prompt 6
I asked 50
RESULT not optimal

Once the prompt runs we get output which we save as tab separated format.
Some examples of the output from the above prompts (out of total 50):
Does the model account for the increased default risk associated with longer-term auto loans?
How does the model evaluate borrower credit and liquidity constraints in relation to loan term choices?
Is the model designed to detect higher-risk borrowers who may self-select into longer-term loans?
Does the model consider the impact of lender experience with longer-term loans on default rates?
These are the questions that will be used by the API and passed on the backend which we will show in section 4.

IV. PROPOSED FULL STACK FRAMEWORK

We propose a full stack framework to integrate the questions collected from the public data using an API to query Bank's proprietary models trained in logistics regression.

The front end is based on public data while the backend is proprietary models of the bank. We have proposed different methodologies to use the relevant questions and curate and run based on current frameworks to address some of the ethical challenges that are pointed out in recent years. We propose the API to convert the cleaned questions and then pass it on as a query to the backend to run on prosperity models.
In figure 4 we have showed the frontend, backend and connections to define public and private spaces.

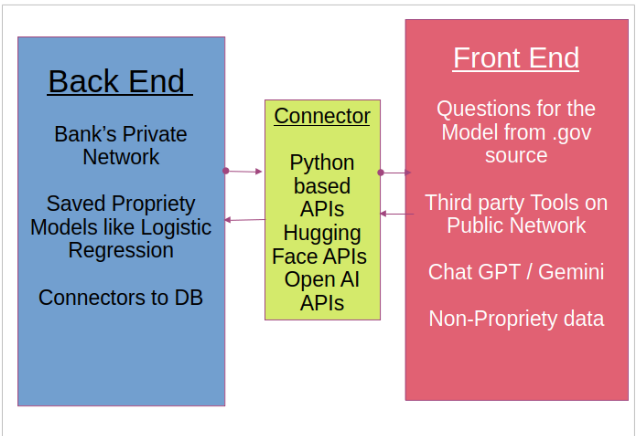
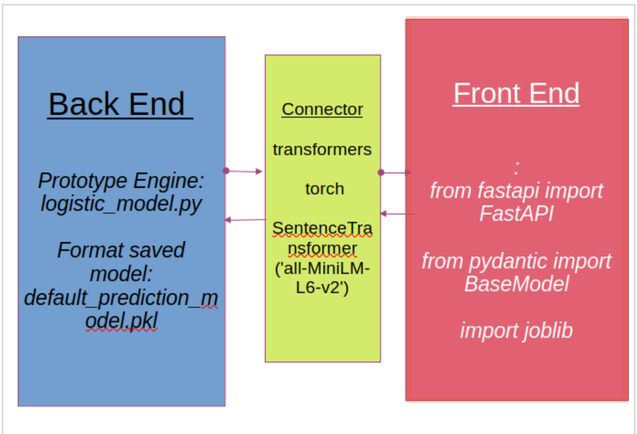


Figure 4: Outline of Proposed Full Stack Framework



Figures 5: Libraries Used in the Proposed Framework

In figure 5, we have shown the python repositories used in the prototype. In the below code snippet we have shown how we can use a pre trained sentence transformer using the hugging face library.
from transformers import AutoTokenizer, AutoModelForSequenceClassification
import torch
from sentence_transformers import SentenceTransformer, util
# Load a pre trained sentence transformer model for semantic matching
model = SentenceTransformer('all-MiniLM-L6-v2')
In the table 2 shown below, we tested 10 variantes and prompts and found 90% accuracy for the all-MiniLM-L6_v2[13] pre-trained model.
The versions used to generate results in this work is as follows:
tensorboard==2.18.0

sentence-transformers==3.3.1
transformers==4.46.3
all-MiniLM-L6_v2

We gave simple prompts and the results are shown in Table 2 but have shown the viability of the framework goes much beyond this work. For example, re-running the base mode saved model as MODEL_PATH = "default_prediction_model.pkl"

Table 2: Backend Results

Prompt Variants	Accuracy	Sample
Re-run model	90%	10
Re-train model	90%	10
Update with new inputs	90%	10

V. CONCLUSION

Gen AI model for generating regulatory questions which can then be linked to a pre-trained API model for full spectrum model testing, validation and review is implemented on a small scale. The questions were fetched using publicly available LLM tools like ChatGPT. Results of accuracy of query generation of models are shown which are ranked by three analysts (volunteers). We also found that publicly available platforms require multiple prompts as they want to keep the outputs computationally inexpensive and also relevant so they do not follow our instructions often and require repetition of prompts. Full stack model to develop end to end automated testing models is proposed which can handle various variations of the questions and then map them to the operation. The expansion and future work will be to link regulatory questions and get the answers using the newer model of Open AI. The proposed framework can enhance the productivity of analysts and enhance accuracy of risk models. Future work can include increasing the scale of queries, addressing the technical challenges of adoption, increasing to a larger pool of analysts, enhancing and adding more metrics.

REFERENCES

[1] A. Balakrishnan, "Leveraging artificial intelligence for enhancing regulatory compliance in the financial sector," Social Science Research Network, 2024. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4842699

[2] O.A. Bello, "Machine learning algorithms for credit risk assessment: An economic and financial analysis," International Journal of Management Technology, vol. 10, no. 1, Dec. 2023. Available from: https://doi.org/10.37745/ijmt.2013

[3] C. Challoumis, "What challenges does AI present to the cycle of money and economocracy?" SSRN Electronic Journal, 2024. Available from: https://doi.org/10.2139/ssrn.4975035

[4] X. Cheng et al., "Combating emerging financial risks in the big data era: A perspective review," Fundamental Research, vol. 1, no. 5, pp. 595–606, Sep. 2021. Available from: https://doi.org/10.1016/j.fmre.2021.08.017

[5] W. of Conferences, "XVI international scientific conference," Tallinn, Estonia, 17-18 Oct. 2024. Available from: https://doi.org/10.1109/APEIE59731.2023

[6] M. Doumpos et al., "Operational research and artificial intelligence methods in banking," European Journal of Operational Research, vol. 306, no. 1, pp. 1–16, Apr. 2023. Available from: https://doi.org/10.1016/j.ejor.2022.04.027

[7] I. Goldstein et al., "Big data in finance," The Review of Financial Studies, vol. 34, no. 7, pp. 3213–3225, Jun. 2021. Available from: https://doi.org/10.1093/rfs/hhab038

[8] H. Herrmann and B. Masawi, "Three and a half decades of artificial intelligence in banking, financial services, and insurance: A systematic evolutionary review," Strategic Change, vol. 31, no. 6, pp. 549–569, 2022. Available from: https://doi.org/10.1002/jsc.2525

[9] D.K. Nguyen et al., "Big data, artificial intelligence and machine learning: A transformative symbiosis in favour of financial technology," European Financial Management, vol. 29, no. 2, pp. 517–548, Mar. 2023. Available from: https://doi.org/10.1111/eufm.12365

[10] H. Sadok et al., "Artificial intelligence and bank credit analysis: A review," Cogent Economics & Finance, vol. 10, no. 1, pp. 2023262, Dec. 2022. Available from: https://doi.org/10.1080/23322039.2021.2023262

[11] Y. Wang et al., "Can fintech improve the efficiency of commercial banks?—An analysis based on big data," Research in International Business and Finance, vol. 55, p. 101338, Jan. 2021. Available from: https://doi.org/10.1016/j.ribaf.2020.101338

[12] K. Yu et al., "A deep reinforcement learning approach to enhancing liquidity in the U.S. municipal bond market: An intelligent agent-based trading system," International Journal of Innovative Research in Engineering and Management, vol. 11, no. 6, pp. 43–54, Dec. 2024. Available from: https://doi.org/10.5281/zenodo.14184756

[13] Hugging Face, "all-MiniLM-L6-v2: A lightweight model for generating sentence embeddings," Hugging Face, Oct. 2020. DOI: 10.5281/zenodo.4633956. Available from: https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

ABOUT THE AUTHORS

Satyadhar Joshi is currently working with Bank of America, Jersey City, NJ as Assistant Vice President in Global Risks and Analyst Department.