Financial and Corporate Risk Analysis Using Large Language Models (LLMs)

Dr. (Mrs.) Sujata Khedkar[#1], Srushti Satish Sambare[#2], Tasmiya Sarfaraz Khan[#3],

Purtee Santosh Mahajan[#4], Ketaki Dhananjay Nalawade[#5]

[#]*Computer Engineering Department, Vivekanand Education Society's Institute of Technology, Chembur, Mumbai - 400071, India*

[1]sujata.khedkar@ves.ac.in

[2]2021.srushti.sambare@ves.ac.in

[3]2021.tasmiya.khan@ves.ac.in

[4]2021.purtee.mahajan@ves.ac.in

[5]2021.ketaki.nalawade@ves.ac.in

Abstract— **The purpose of this paper is to examine how Large Language Models (LLMs), specifically Gemini, may be used in financial and corporate risk analysis. In particular, we utilize LLMs to perform risk assessment activities using corporate data in the form of Corporate Earnings Calls, Annual Public Reports and Environmental, Social, and Governance (ESG) reports. Our objective is to understand the potential risks of doing business by undertaking individual risk assessment in these data sources. All these risk assessments are subsequently combined into one risk report, reflecting all possible risk exposures associated with the company's activities. This technique improves the overall efficiency of risk assessment due to the decrease of time needed for its completion while increasing the quality of available information for decision-making.**

Keywords— Large Language Models (LLMs), Gemini, Financial Risk Analysis, Corporate Risk Assessment, Corporate Earnings Calls, Annual Public Reports, ESG Reports, Generative AI, Risk Reporting, Corporate Governance

I. Introduction

In an era marked by increasing economic globalization, businesses present tremendous opportunities to expand but also face increasing pressures from competitive markets [1]. As organizations seek to capture this environment actively, risk management and compliance have become key components of corporate governance, especially in financial services ensuring that companies can maintain trust and stability [2].

Today's complex corporate environment calls for a comprehensive risk management strategy to build organizational resilience and sustainability [3]. Financial pitfalls, in particular, are receiving accelerating engrossment due to their potentially eloquent impact on the organization's future. These risks can arise from a variety of sources, such as market fluctuations, regulatory changes, or faulty internal controls, making it necessary for companies to take a holistic approach to risk assessment and management [4].

The field of risk assessment and management has been firmly established over the last 30–40 years, during which time it has evolved into an organized scientific discipline [5]. The decision-makers need to see beyond the risk evaluation; they need to combine the risk information they have received with information from other sources and on other topics [5].

The objective of this paper is to explore Large Language Models (LLMs) like Google's Gemini, which can enhance the risk assessment and management process by analyzing unstructured data from sources such as corporate earnings calls, annual public reports, and ESG reports. These insights, when combined, provide a solid foundation for more accurate and comprehensive risk assessments, ultimately contributing to better corporate decision-making.

II. Related work

The exploration of financial risk prediction and management using advanced machine learning models has been expanding, offering innovative techniques for understanding and mitigating financial risks.

In the paper RiskLabs: Predicting Financial Risk Using Large Language Models Based on Multi-Sources Data [6], the authors present the RiskLab framework, which integrates multiple modules to handle diverse sources of information such as earnings conference calls and time-series data. By employing techniques like self-attention and Bayes-Value at Risk (VaR) forecasting, the framework demonstrates a sophisticated approach to combining news filtering and contextual compression to achieve more flexible training. A major merit of this framework is its ability to generate nuanced, multi-task predictions that provide investors with comprehensive insights into market conditions. However, a significant limitation is the reliance on large language models (LLMs), which can sometimes produce inaccurate or misleading responses, such as hallucinations, and struggle with real-time market updates due to outdated data sources.

In Enhancing Credit Risk Reports Generation Using LLMs: An Integration of Bayesian Networks and Labeled Guide Prompting [7], the authors focus on generating high-quality credit risk reports through a combination of Labeled Guide Prompting (LGP) and Bayesian networks. This methodology ensures the outputs from GPT-4 align closely with the specific requirements of credit risk analysis. A notable merit is that the generated reports were statistically preferred over traditional human-generated reports. The integration of LLMs into the credit risk process improves efficiency and scalability, making it an attractive solution for large-scale financial assessments. However, GPT-4's tendency to equally weigh all features poses a limitation, potentially failing to account for the importance of various risk factors in real-world assessments. Additionally, the model's occasional inaccuracies raise concerns about the reliability of its output.

The third paper, Fusing LLMs and Knowledge Graphs for Formal Causal Reasoning Behind Financial Risk Contagion [8], offers a novel approach to understanding how financial risks spread across interconnected entities. The Risk Contagion Causal Reasoning Model integrates Financial Knowledge Graphs (KGs) with LLMs to provide a formal causal understanding of financial contagion. By employing a fusion module and Sankey diagrams, the authors successfully show the pathways through which risks propagate. This model enhances the ability to develop targeted strategies for preventing financial crises, identifying critical nodes in the risk network. However, the challenge lies in integrating unstructured language data with structured graph data, and the model's success is dependent on the quality of the underlying financial KGs, which may not always be available or accurate.

Finally, Deep Learning Model-Driven Financial Risk Prediction and Analysis [9] explores the application of generative deep learning models for simulating financial time series data and predicting Value at Risk (VaR). By comparing the performance of different generative models, such as GANs (Generative Adversarial Networks), the paper highlights significant improvements in VaR prediction accuracy. This approach is particularly effective at capturing the complex distribution patterns in financial data. However, a notable limitation is the model’s inability to quickly adapt to sudden market changes, which can lead to inaccurate predictions during crises, where rapid responses are critical.

III. Input Sources

The system deals with three input sources. They are elaborated as follows:

A. Corporate Earnings Call

Earnings calls play a significant role in increasing investment in the company by providing economic communication. Earnings conference calls are held following the quarterly release of a company's earnings and have grown in popularity in recent years, owing to their ease of access via modern communication media (e.g., programs like EarningsCast, interactive investor-relation websites)[10][11]. The goal of these calls is to inform the market about the firm's future strategy and tactics, as well as remark on the previous quarter's revenue streams and costs [10][11].

B. Annual Public Report

An annual public report is a detailed report that public companies are required to present annually to their shareholders and the local tax office which can prepare auditor’s reports. Annual report of a company contains the directors’ report, the auditor’s report, the financial statements and the schedules and notes to the accounts [12]. For all the companies under the Ministry of Corporate Affairs (under Section 217 of Companies Act, 2013), (approx. 2263265 companies), it is mandatory to publish an annual report.

C. Environmental, Social and Governance Report (ESG)

ESG reporting is the disclosure of environmental, social and corporate governance data. As with all disclosures, its purpose is to shed light on a company’s ESG activities while improving investor transparency and inspiring other organizations to do the same. Since ESG reports summarize the qualitative and quantitative benefits of a company’s ESG activities, investors can screen investments, align investments to their values, and avoid companies with the risk of environmental damage, social missteps or corruption [13].

IV. Proposed solution

The user can upload either one or two or all of the three input sources i.e. corporate earning call, annual public report, ESG report. A risk assessment will be done on each of these sources. Risk analysis will be done based on the following parameters:

1. Highlight the statements depicting risk
2. Highlight the statements with negative impact
3. For numerical data, the system will distinguish between positive and negative trends. For example, a decrease in profits by x% will be classified as negative, whereas a decrease in losses by x% will be considered as positive. This nuanced approach will ensure that negative statements are accurately identified while recognizing any improvements.

The assessment results are further analyzed and suggestions to lower the risk are given. Finally, these assessment results would be combined to form a risk report.

V. METHODOLOGY

Step 1: Uploading Files

1. The document is given as input using steamlit's file uploader component.
2. The user clicks on the 'Process PDF' button.

Step 2: Processing the Files

1. The process_pdf() function is called. This function reads the text content from each page of the PDF file using the 'PdfReader' from the PyPDF2 library.
2. The text content from each page is concatenated to form a single string representing the entire document.
3. The get_text_chunks() function is called. This function takes the text string as input and splits it into smaller chunks. It uses 'RecursiveCharacterTextSplitter' from the LangChain Library to split the text into chunks of size 5000 with an overlapping of 1000.
4. The get_vector_store() function is called. In the context of Natural Language Processing, the Vector store is a data structure which stores the vector representation of textual data. This function takes the text chunk as input. It generates vector embeddings for each text chunk using the Google Generative AI embeddings..These embeddings are stored in a vector store created using the FAISS Library.
5. The get_conversational_chain() function is called. This function creates a conversational chain using the LangChain library. It takes the vector store and a prompt template as input.
6. First, it initializes the gemini-1.5-flash model which is our large language model for generating responses.
7. It creates a memory component 'ConversationBufferMemory' to store the conversation history.

Step 3: Handling User Queries

1. Now as the PDFs are processed, the user enters a question and clicks on get response.
2. The user_input function is called. It takes the user's question as input. This function interacts with the conversational chain stored in the streamlit session state to retrieve a response.
3. The vector representation of the query is created and is compared with the vector representations present in the vector store.
4. The most similar vectors are then retrieved based on the similarity.
5. The conversation history is updated and displayed to the user.

Step 4: Displaying the Response

1. First, the code checks if there is any conversation history stored in streamlit session state. If it is there, it iterates through each message in the history.
2. For each message, it checks if the message index ('i') is even or odd. If it is even, it is the user's message and if it is odd, it is the bot's message.
3. If it is the response, it checks if the response contains table formatting by looking for '|' and '---' characters.
4. If yes, then it parses the response into a dataframe, eliminates dirty values and displays it as a table.
5. Then, it extracts the data from the table into a data.csv file.
6. This data is string data. It contains words like million, billion, M, B, %, $, etc. We convert those characters into empty string and then the entire value into float and plot the graph.
7. But, if there is no table in the response, it is displayed as a plain text.

Step 5: Generating FAQs

1. We have predefined categories and questions.
2. The response is retrieved the same way as user queries are answers.

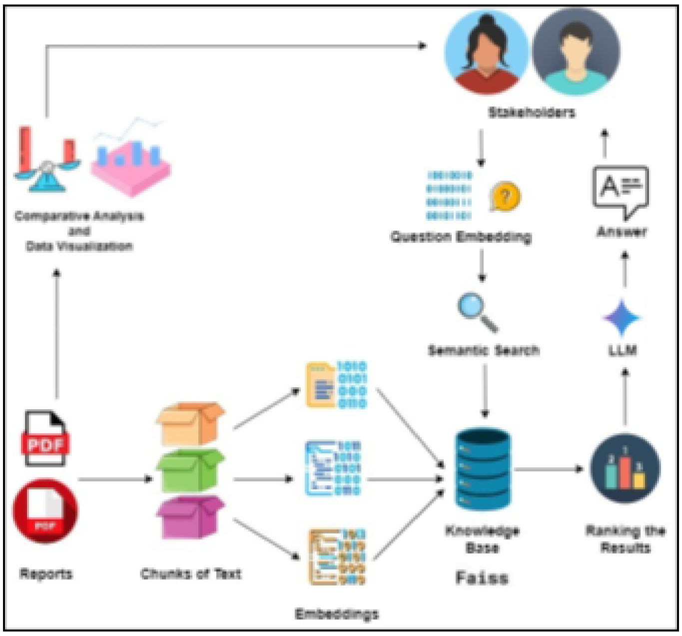

Fig. 1 System Architecture based

VI. RESULTS

The input taken here is a sample document which contains some information about the company XYZ Corporation.

XYZ Corporation Financial Overview

In the fiscal year ending December 31, 2023, XYZ Corporation reported a total revenue of $1.2 billion, reflecting a modest increase of 5% compared to the previous year. However, the company's net profit experienced a significant decline, dropping by 15% to $150 million. This downturn is attributed to rising operational costs and increased competition in the market.

XYZ Corporation has implemented several strategic initiatives to address these challenges. The company has invested heavily in technology upgrades, allocating $100 million towards enhancing its digital infrastructure and improving operational efficiencies. This strategic pivot aims to streamline processes and reduce costs in the long run.

Despite these efforts, the company acknowledges the inherent risks associated with its expansion strategy. The volatility in raw material prices poses a threat to profit margins, as indicated by a 10% increase in material costs over the past year. Additionally, the ongoing geopolitical tensions in key markets have raised concerns regarding supply chain stability, potentially impacting product availability and pricing.

Looking forward, XYZ Corporation remains committed to diversifying its product offerings and exploring new market opportunities. The management team is optimistic about achieving a revenue target of $1.5 billion by the end of 2024. However, they also caution that achieving this goal will require navigating significant market uncertainties, including regulatory changes and economic fluctuations.

In summary, while XYZ Corporation has laid out a comprehensive strategy to enhance profitability and growth, the company remains vigilant of the risks that could affect its financial health and operational success.

Fig. 2 Input Document - Financial Overview of XYZ Corporation

The proposed solution will yield the following outcomes:

A. Statements Depicting Risk

The statements depicting risk are the excerpts or phrases that highlight the potential risks to the concerned company. These risks can be regarding the operations, finances, reputation, etc. The challenges depicting risk can be regulatory changes, supply chain disruptions, economic downturns, cybersecurity threats, or competitive pressures.

The Gemini-1.5-flash model identifies them using the following:

1) *Semantic Understanding*

Using natural language understanding (NLU), Gemini identifies sentences and phrases that convey uncertainty, challenges, or potential threats.

2) *Keyword Detection*

It scans for risk-related terms such as "challenge," "uncertainty," "adverse impact," "vulnerability," and "risk."

3) *Contextual Analysis*

By analyzing the surrounding context, Gemini differentiates between routine business updates and genuine risk indicators.

4) *Training Data*

The model is fine-tuned on large datasets containing risk statements from similar financial reports, enabling it to recognize patterns.

Fig. 3 showcases the statements depicting risks that were extracted by the system from the financial overview which was given as an input.

Statements Depicting Risk:

- "The company's net profit experienced a significant decline, dropping by 15% to $150 million."
- "This downturn is attributed to rising operational costs and increased competition in the market."
- "The volatility in raw material prices poses a threat to profit margins, as indicated by a 10% increase in material costs over the past year."
- "Additionally, the ongoing geopolitical tensions in key markets have raised concerns regarding supply chain stability, potentially impacting product availability and pricing."
- "Despite these efforts, the company acknowledges the inherent risks associated with its expansion strategy."
- "However, they also caution that achieving this goal will require navigating significant market uncertainties, including regulatory changes and economic fluctuations."

Fig. 3 Statements Depicting Risk for XYZ Corporation

B. Negative Impact Statements

The statements that directly mention the adverse effects faced by the company or the effects that the company may face in the future. This could be due to some internal or external factors. The examples can include losses from lawsuits, revenue declines due to market conditions, or negative feedback from stakeholders.

The Gemini-1.5-flash model identifies them using the following:

1) *Sentiment Analysis*

Gemini uses sentiment analysis to classify sections of text as negative, neutral, or positive, flagging phrases that have a distinctly negative tone.

2) *Phrase Extraction*

It looks for phrases indicating setbacks, losses, or failures, such as "declined revenue," "adverse effects," "unexpected losses," or "regulatory penalties."

3) *Section Focus*

Specific sections, like "Management Discussion and Analysis" or "Operational Challenges," are prioritized for deeper analysis.

Fig. 4 showcases the negative impact statements that were extracted by the system from the financial overview which was given as an input.

Negative Impact Statements:

- "The company's net profit experienced a significant decline, dropping by 15% to $150 million."
- "Rising operational costs"
- "Increased competition in the market"
- "10% increase in material costs over the past year"
- "Concerns regarding supply chain stability, potentially impacting product availability and pricing"
- "Significant market uncertainties, including regulatory changes and economic fluctuations"

Fig. 4 Negative Impact Statements for XYZ Corporation

C. Quantifiable Data Classified Into Positive And Negative Trends

This involves analyzing numerical data provided in the reports, such as financial statements, to determine trends. Positive trends include growth in revenue, increased profit margins, or higher customer acquisition rates, while negative trends include declining sales, rising debts, or shrinking market share.

The Gemini-1.5-flash model identifies them using the following:

1) *Data Extraction*

Gemini identifies and extracts numerical data from financial tables, charts, and statements, including income statements, balance sheets, and cash flow reports.

2) *Trend Detection*

Using historical comparisons (e.g., year-over-year or quarter-over-quarter changes), it determines whether trends are positive or negative.

3) *Sentiment and Impact Analysis*

Phrases accompanying numerical data, such as "strong growth" or "significant decline," are analyzed to classify trends effectively.

4) *Rule-Based Models*

Gemini applies predefined financial rules (e.g., a decline in net income over consecutive periods is flagged as negative) to supplement its ML-based approach.

Fig. 5 showcases the quantifiable data classified into positive and negative trends that were extracted by the system from the financial overview which was given as an input.

- **Negative:** Net profit decreased by 15% to $150 million.
- **Negative:** Material costs increased by 10% over the past year.
- **Positive:** Revenue increased by 5% to $1.2 billion.
- **Positive:** Investment in technology upgrades: $100 million.
- **Target:** Revenue target of $1.5 billion by the end of 2024.

Fig. 5 Quantifiable Data Classified into Positive and Negative Trends for XYZ Corporation

D. An Analysis of All The Extracted Data - Risk Analysis

This step involves synthesizing all identified risks, impacts, and trends into a coherent narrative to assess the overall risk profile of the company. It includes understanding the likelihood of risks materializing, their potential impact, and the company's preparedness to address them.

The Gemini-1.5-flash model gives the assessment summary using the following:

1) *Correlation of Insights*

Gemini synthesizes extracted data, connecting risk statements, negative impacts, and quantifiable trends to create a comprehensive picture of the company's risk profile.

2) *Model Outputs*

It uses probabilistic models to evaluate the likelihood and severity of risks.

3) *Risk Taxonomy*

Gemini maps extracted data to predefined risk categories (e.g., operational, financial, or compliance risks) based on its understanding of similar documents.

Fig. 6 showcases the risk assessment summary which was generated by the system from the financial overview which was given as an input and the previous results.

Assessment Summary:

XYZ Corporation faces several significant risks that could impact its financial health and operational success.

- **Profitability Decline:** The company experienced a substantial decline in net profit, attributed to rising operational costs and increased competition. This trend could continue if not addressed effectively.
- **Raw Material Price Volatility:** Fluctuations in raw material prices pose a threat to profit margins, as evidenced by the 10% increase in material costs. This could further impact profitability if not mitigated.
- **Supply Chain Disruptions:** Geopolitical tensions and potential supply chain disruptions could negatively impact product availability and pricing, affecting both revenue and customer satisfaction.
- **Market Uncertainties:** Regulatory changes and economic fluctuations add to the existing challenges, requiring careful navigation and strategic adjustments to achieve the targeted revenue growth.

Fig. 6 Risk Assessment for XYZ Corporation

E. Risk Mitigation Suggestions

Based on the risk analysis, this involves proposing actionable steps the company can take to reduce exposure to identified risks. Suggestions can be operational, strategic, or technological and may include diversifying supply chains, enhancing cybersecurity measures, or adopting new technologies.

The Gemini-1.5-flash model gives the suggestions using the following:

1) *Generative Capabilities*

Leveraging its generative abilities, Gemini provides actionable recommendations based on best practices and historical data.

2) *Trend-Based Recommendations*

For example, if a downward trend in revenue is identified, Gemini might suggest cost-cutting measures or diversifying revenue streams.

3) *Mitigation Mapping*

The model correlates risks to solutions provided in the company report or external sources, enabling it to suggest precise mitigation strategies.

Fig. 7 showcases the risk mitigation suggestions which were generated by the system from the financial overview which was given as an input.

Risk Mitigation Suggestions:

- **Cost Optimization:** Implement cost reduction strategies across all operations to address the rising operational costs. This could include streamlining processes, negotiating better deals with suppliers, and exploring alternative sourcing options.
- **Competitive Differentiation:** Develop a strong competitive strategy to address the increasing market competition. This could involve product innovation, targeted marketing campaigns, and enhancing customer service.
- **Raw Material Hedging:** Explore hedging strategies to mitigate the impact of raw material price volatility. This could involve using forward contracts or other financial instruments to lock in prices.
- **Supply Chain Diversification:** Diversify supply chains to reduce reliance on specific regions and minimize the impact of geopolitical tensions. This could involve sourcing materials from multiple locations or establishing alternative manufacturing facilities.
- **Strategic Market Analysis:** Continuously monitor regulatory changes and economic fluctuations to proactively adapt the business strategy. This could involve conducting regular market research, engaging with industry experts, and developing contingency plans.

By actively addressing these risks and implementing mitigation strategies, XYZ Corporation can improve its resilience and enhance its chances of achieving its ambitious growth targets.

Fig. 7 Risk Mitigation for XYZ Corporation

VII. Technologies Used

A. Google Gemini API

Gemini is a family of highly capable multimodal models developed at Google [14]. The system is powered by the Gemini-1.5-flash model.

B. FAISS (Facebook AI Similarity Search)

Faiss is a library for ANNS. The core library is a collection of source files written in standard C++ without dependencies. Faiss is used in many configurations. Hundreds of vector search applications rely on it, both within Meta. [15] and externally. The proposed system deals with a large amount of textual data. Hence, FAISS is always preferred for scalability as it avoids a drop in performance. The purpose of this utilization is to perform similarity search on the vectorized representations of the documents and for a quick and easy retrieval.

C. Streamlit

Streamlit is an open-source Python framework for data scientists and AI/ML engineers to deliver dynamic data apps with only a few lines of code. [16] With this framework, you can easily build interactive visualization plots, models, and dashboards without having to worry about the underlying web framework or deployment infrastructure used in the backend [17]. Even though a Streamlit app is easy to build and deploy, it is not scalable. Hence, for large scale applications, frameworks like Flask, a Python-based framework would be preferred.

D. Langchain

Langchain is a custom Large Language Model tailored for organizations [18]. LLMs have been rapidly adopted due to their capabilities in a range of tasks, including essay composition, code writing, explanation, and debugging [19]. They accept a text string (prompt) and output a text string [19].

VIII. Future Scope

The current implementation of the risk assessment system has considerable utility at accelerating evaluation of the financial documents and risk reports' compilation. However, there are multiple areas where the system can be extended and improved to better meet the needs of organizations and enhance its capabilities:

A. Customizable Report Formats

Another important feature which should be anticipated in the future is the possibility of an implementation to define the layout and structure of the risk reports produced. It has become a norm for different companies to have preferred style of reporting, including sections preferred in the report, preferred kind of diagrams or tables, or industry-specific language. It was proposed that, through a reporting feature that can be customized, the user would be able to prepare the report according to internal rules or other requirements.

B. Linking of documents to more than one document type

The present system of analyzing and interpreting the PDF-file format of financial reports may be expanded in subsequent versions for other formats, such as Excel tables, Word documents or even access to database information. This would enhance flexibility of the system in handling accounting information from various data feeds from different sources, and would also ensure compatibility to various documentation processes by companies.

IX. Conclusion

This research outlines a simple way of automating financial risk assessment through Gemini 1.5 Flash coupled with enhanced user interface. The system is extremely effective to scan financial documents, flag risks and offer actionable intelligence. Automating risk analysis therefore enables organizations to take quicker and better decisions. Additional additions such as report features, or industry-specific models will enhance the system's flexibility making the system relevant for companies as they grapple with sophisticated risk levels.

References

[1] Environmental and Public Health, Journal of. "Retracted: Prediction and Analysis of Corporate Financial Risk Assessment Using Logistic Regression Algorithm in Multiple Uncertainty Environment." *Journal of Environmental and Public Health 2023 (2023): n. pag.*

[2] Judijanto, Loso, Sitti Hartati Hairuddin, Subhan Subhan and Baren Sipayung. "Analysis of the Effect of Risk Management and Compliance Practices on Financial Performance and Corporate Reputation in the Financial Industry in Indonesia." *The Es Accounting And Finance (2024): n. pag.*

[3] Balaji, Saradha & Shreshta, Lolakpuri & Sujatha, K.. (2024). A Study on Risk Management in Corporate Business. Involvement International Journal of Business. 1. 197-209. 10.62569/iijb.v1i3.26.

[4] Sun, Weihua. (2024). Research on Corporate Financial Risk Management and Countermeasures. International Journal of Global Economics and Management. 3. 59-65. 10.62051/IJGEM.v3n1.07.

[5] Aven, T. (2015). Risk assessment and risk management: Review of recent advances on their foundation. European Journal of Operational Research, 253(1), 1–13. https://doi.org/10.1016/j.ejor.2015.12.023

[6] Cao Y, Chen Z, Pei Q, Dimino F, Ausiello L, Kumar P, Subbalakshmi KP, Ndiaye PM. RiskLabs: Predicting Financial Risk Using Large Language Model Based on Multi-Sources Data. arXiv preprint arXiv:2404.07452. 2024 Apr 11.

[7] Teixeira, Ana & Marar, Vaishali & Yazdanpanah, Hamed & Pezente, Aline & Ghassemi, Mohammad. (2023). Enhancing Credit Risk Reports Generation using LLMs: An Integration of Bayesian Networks and Labeled Guide Prompting. 340-348. 10.1145/3604237.3626902.

[8] Yu G, Wang X, Li Q, Zhao Y. Fusing LLMs and KGs for Formal Causal Reasoning behind Financial Risk Contagion. arXiv preprint arXiv:2407.17190. 2024 Jul 24.

[9] Yang, Tianyi & Li, Ang & Xu, Jiahao & Su, Guangze & Wang, Jiufan. (2024). Deep Learning Model-Driven Financial Risk Prediction and Analysis. 10.20944/preprints202406.2069.v1

[10] Vivien E. Jancenelle Susan Storrud-barnes Rajshekhar Javalgi , (2017)," Corporate Entrepreneurship and Market Performance: A Content Analysis of Earnings Conference Calls ", Management Research Review, Vol. 40 Iss 3 pp. -

[11] Kimbrough, M.D. (2005), "The effect of conference calls on analyst and market underreaction to earnings announcements", The Accounting Review, Vol. 80 No. 1, pp. 189-219.

[12] P. K. Aithal, D. A. U. and G. M., "Analyzing Tone of the Annual Report - An Indian Context," 2022 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), THIRUVANANTHAPURAM, India, 2022, pp. 536-542, doi: 10.1109/SPICES52834.2022.9774247. keywords: {Instruments;Signal processing algorithms;Companies;Signal processing;SPICE;Stock markets;Information technology;Portfolio Management;Tone;Initial Public Offerings;Parallel Algorithms;Message Passing Interface}

[13] Tocchini, F., & Cafagna, G. (2022, March 9). The ABCs of ESG reporting: What are ESG and sustainability reports, why are they important, and what do CFOs need to know. https://www.wolterskluwer.com/en/expert-insights/the-abcs-of-esg-reporting#:~:text=ESG%20reporting%20is%20the%20disclosure,organizations%20to%20do%20the%20same.

[14] G. Team *et al.*, "Gemini: a family of highly capable multimodal models," *arXiv.org*, Dec. 19, 2023. https://arxiv.org/abs/2312.11805

[15] M. Douze *et al.*, "The Faiss library," *arXiv.org*, Jan. 16, 2024. https://arxiv.org/abs/2401.08281

[16] "Streamlit Docs." https://docs.streamlit.io/

[17] Sreeram a, Adith & Sai, Jithendra. (2023). An Effective Query System Using LLMs and LangChain. International Journal of Engineering and Technical Research. 12.

[18] K. Pandya and M. Holia, "Automating Customer Service using LangChain: Building custom open-source GPT Chatbot for organizations," *arXiv.org*, Oct. 09, 2023. https://arxiv.org/abs/2310.05421

[19] Topsakal, Oguzhan & Akinci, T. Cetin. (2023). Creating Large Language Model Applications Utilizing LangChain: A Primer on Developing LLM Apps Fast. International Conference on Applied Engineering and Natural Sciences. 1. 1050-1056. 10.59287/icaens.1127.