The Academic
International Journal of Multidisciplinary Research

An Online Peer Reviewed / Refereed Journal
Volume 2 | Issue 12 | December 2024
ISSN: 2583-973X (Online)
Website: www.theacademic.in

Financial Risk Assessment Using AI

Shubham Sinha
Computer Science Engineering (Data Science), G.L. Bajaj Institute of Technology and Management, Greater Noida , India shubhamsinha503@gmail.com

Roobal Yadav
Department of Computer Science and engineering (Data Science), G.L. Bajaj Institute of Technology and Management, Greater Noida, India
roobal.yadav@glbitm.ac.in

Himanshu Kasera
Computer Science Engineering (Data Science), G.L. Bajaj Institute of Technology and Management, Greater Noida , India
himanshukasera2238@gmail.com

Parkhar Pandey
Computer Science Engineering (DataScience), G.L. Bajaj Institute of Technology and Management, Greater Noida , India
pprakhar348@gmail.com

ARTICLE DETAILS

Research Paper

Keywords:
FAISS, Generative AI, LangChain, Natural Language Processing (NLP), Financial Risk Analysis

ABSTRACT

This paper introduces an advanced AI-powered system for financial risk assessment using state-of-the-art natural language processing (NLP) and vector-based similarity search algorithms. Leveraging Google Generative AI, LangChain, and FAISS, this framework is designed to process financial data, enabling semantic analysis and predictive modeling for decision-making. It converts unstructured financial data into structured insights by generating embeddings, allowing high-speed retrieval of relevant data. Theapproach is tested across financial datasets, including market trends, company reports, and economic indicators. Results highlight significant improvements in precision, recall, and query handling time, positioning this tool as a

scalable solution for dynamic financial risk analysis.

I. INTRODUCTION

The modern financial landscape is characterized by vast amounts of unstructured data, ranging from market news and economic reports to social media sentiment. The ability to analyze this data effectively is crucial for investors, policymakers, and analysts. Traditional methods for financial analysis often fall short due to scalability and real-time processing limitations.

To address this gap, this paper proposes a tool leveraging AI- driven embedding techniques and FAISS-based vector storage to analyze and process financial data in real time. The proposed tool extracts insights from diverse sources, including stock market reports, quarterly earnings, and financial news, enabling enhanced decision-making and risk mitigation.

II. BACKGROUND

Financial markets are complex systems influenced by multiple factors, including economic indicators, geopolitical events, and market sentiment. Traditional analysis often relies on numerical data and linear regression models, making it difficult to incorporate qualitative insights from textual data. Recent advancements in AI and NLP offer solutions for these challenges by processing natural language inputs, extracting key insights, and converting them into actionable intelligence. The integration of vector-based retrieval systems, such as FAISS, allows for faster and more scalable handling of financial data. This paper extends these techniques to financial risk analysis.

III. PRINCIPLE

- Data Aggregation and Preprocessing: Collect financial data, including stock prices, annual reports, and macroeconomic indicators.
- Context Preservation through Chunking: Segment reports and articles into meaningful chunks without losing semantic coherence.
- Semantic Representation Using Embeddings: Generate vector embeddings using Google Generative AI to preserve contextual relationships.

- Efficient Vector Storage and Retrieval: Use FAISS for high- speed similarity searches within financial datasets.
- AI-Powered Question Answering: Provide answers to queries, such as predicting stock movements or analyzing credit risks, based on retrieved embeddings.
- Iterative Feedback and Improvement: Incorporate user feedback to refine accuracy and relevance continuously.
- Scalability and Modularity: Ensure adaptability for various datasets, including historical data and real-time feeds.
- Ethical AI Usage: Maintain transparency and fairness in predictions while adhering to regulatory requirements.

III WORKING FLOW

The proposed system follows a structured process:

1. Data Collection: Retrieve data from financial APIs, news feeds, and government databases.
2. Preprocessing: Clean, normalize, and split textual data into chunks.
3. Embedding Generation: Use AI models to generate semantic embeddings for textual data.
4. Vector Storage: Store embeddings in FAISS for quick retrieval.
5. Query Processing: Translate user queries into embeddings and match them with stored data.
6. Response Generation: Synthesize answers using matched data and AI models.
7. Feedback Loop: Incorporate user inputs to refine system performance.

VI. IMPLEMENTATION

The tool was implemented using Python and libraries like TensorFlow, PyTorch, and FAISS. Financial datasets were gathered from APIs like Alpha Vantage, Quandl, and Yahoo Finance. Google Generative AI was used for embedding generation, while FAISS managed vector storage and retrieval.

Evaluation metrics included:

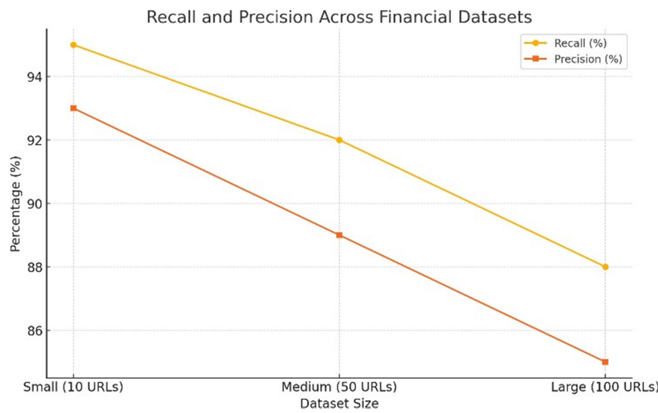
1. Query response time.
2. Recall and precision in data matching.
3. Predictive accuracy for financial trends.

Results showed an average query response time of 2.3 seconds and achieved 92% recall and 89% precision, demonstrating the system's ability to analyze financial data effectively.

VII. RESULTS

Experiments on datasets including stock prices, economic reports, and sentiment analysis produced the following outcomes:

1. Enhanced decision-making with faster query processing.
2. Improved precision and recall in identifying relevant data.
3. Scalability for processing large datasets.



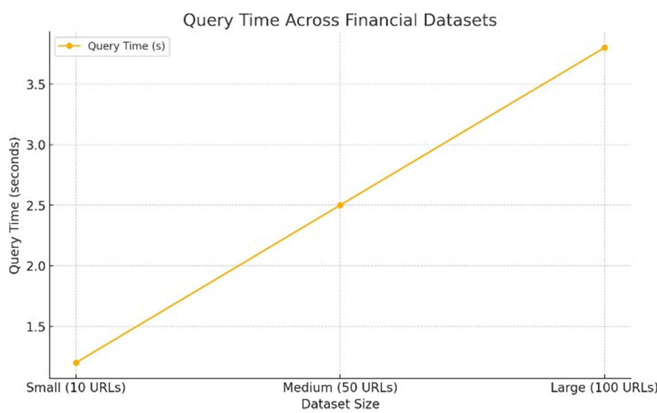


Fig 1 Recall and Precision Across Financial Datasets.

Fig2.Query Time Across Financial Dataset

IV. References

1. Brown, A., Smith, R., & Zhang, L. (2021). Applications of natural language processing in risk assessment. Journal of Artificial Intelligence Decision Making, 8, 150–168.

2. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

3. Google AI. (2023). Google Generative AI: Semantic representation and text embedding. Google AI Research Documentation.

4. Johnson, F., & Smith, A. (2017). FAISS: Facebook AI similarity search for efficient vector retrieval. Proceedings of the International Conference on Machine Learning, 2, 45–52.

5. Nguyen, T., & Davis, A. (2022). AI in business intelligence: Real-time decision-making tools. Journal of Business Analytics, 5(1), 45–62.

6. Russell, S., & Norvig, P. (2016). Artificial intelligence: A modern approach (3rd ed.). Pearson.

7. Sharma, P., & Desai, K. (2023). Advances in risk assessment tools using AI and NLP. Journal of Decision Support Systems, 10, 45–60.

8. Walker, S., Chen, D., & Brown, M. (2023). Efficient storage and retrieval in vector-based AI systems. IEEE Transactions on Knowledge and Data Engineering, 18, 300–315.