

Automated Financial News Analyzer using Large Language Models

Ms. C. Revathi¹

Asst.Professor, Department of AIDS
Annamacharya Institute of
Technology and Sciences, Tirupati –
517520, A.P.

revathiais@gmail.com

K.Mahammad sadak⁴

Student, Department of AIDS
Annamacharya Institute of
Technology and Sciences, Tirupati –
517520, A.P.

Sadakayyan3036@gmail.com

P.Bhuvanewari²

Student, Department of AIDS
Annamacharya Institute of
Technology and Sciences, Tirupati –
517520, A.P.

Bhuvanewari.p0204@gmail.com

A.Hemanth Kumar Reddy⁵

Student, Department of AIDS
Annamacharya Institute of
Technology and Sciences, Tirupati –
517520, A.P.

annemhemanthreddy@gmail.com

Shaik Shajid Basha³

student, Department of AIDS
Annamacharya Institute of
Technology and Sciences, Tirupati –
517520, A.P.

Sajidbasha143@gmail.com

Abstract:

A Market News Analyzer based on Large Language Models (LLMs) is designed to transform unstructured financial news into meaningful insights that support intelligent decision-making. With the rapid growth of digital news sources, investors and analysts face challenges in processing vast amounts of information in real time. This system leverages advanced language models such as GPT-4 and LLaMA to automatically analyze news articles, social media posts, and financial reports. It performs tasks including sentiment detection, entity recognition, and contextual understanding to identify market trends and potential impacts on stocks and sectors. By integrating natural language processing with machine learning techniques, the system provides a scalable and efficient solution for financial data interpretation.

The proposed framework enhances traditional market analysis by offering real-time insights and predictive capabilities. It evaluates the tone and relevance of news content to forecast possible market movements and detect early warning signals for risks. As LLM technology continues to advance, Market News Analyzers are expected to play a crucial role in automated trading systems, portfolio management, and financial forecasting, making them an essential tool in modern financial ecosystems.

Index Keywords: Market News Analysis, Large Language Models (LLMs), GPT-4, LLaMA, Natural Language Processing (NLP), Sentiment Analysis, Financial Text Mining, Stock Market Prediction, Trend Analysis, Risk Assessment, Machine Learning, Artificial Intelligence, Data Analytics, Information Extraction, Real-Time Monitoring

I. INTRODUCTION

The rapid expansion of digital media and online financial platforms has led to an unprecedented surge in the availability of market-related information. News articles, social media posts, analyst reports, and corporate announcements collectively influence investor behavior and market dynamics. However, the sheer volume and velocity of this data make it increasingly difficult for human analysts to process and [3] interpret information effectively. Traditional methods of market analysis rely heavily on manual effort and structured data, which often fail to capture the nuanced insights hidden within unstructured textual content. This challenge highlights the need for intelligent systems capable of automatically analyzing and extracting meaningful information from diverse data sources.

Recent advancements in Natural Language Processing (NLP) and deep learning have paved the way for more sophisticated approaches to text analysis. In particular, Large Language Models such as GPT-4 and LLaMA

have demonstrated remarkable capabilities in understanding context, semantics, and sentiment in human language. These models are trained on vast datasets and can perform complex tasks such as summarization, sentiment [9] [1] classification, and entity recognition with high accuracy. Their ability to process unstructured text in real time makes them highly suitable for financial applications, where timely insights are critical for decision-making.

Market sentiment plays a crucial role in influencing financial markets, often driving price fluctuations beyond what traditional numerical indicators can explain. News related to economic policies, geopolitical events, company performance, and [7] global crises can significantly impact investor confidence and market trends. By leveraging LLMs, it becomes possible to analyze such information at scale and detect patterns that may not be immediately visible. This enables the identification of positive, negative, or neutral sentiments associated with specific stocks, sectors, or the overall market, providing valuable insights for traders and financial institutions..

II. METHODOLOGY

The proposed Market News Analyzer is designed using a multi-stage methodology that systematically processes unstructured financial data and converts it into meaningful insights. The process begins with data acquisition from diverse and reliable sources such as financial news websites, social media platforms, and corporate disclosures. These data streams are collected in real time using APIs and web scraping techniques to ensure continuous updates. The collected raw text is often noisy and unstructured, requiring preprocessing steps such as tokenization, normalization, stop-word removal, and noise filtering to prepare the data for further analysis.

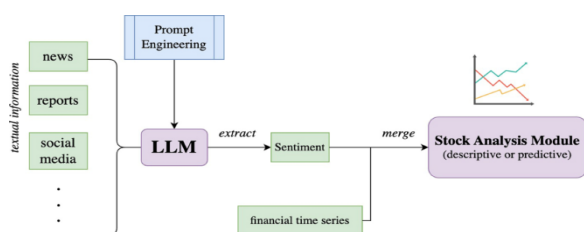


Figure 1: News Sentiment Predictor

Once the data is preprocessed, it is passed into the core processing layer powered by Large Language Models such as GPT-4 and LLaMA. These models perform advanced Natural Language Processing tasks, including sentiment analysis, named entity recognition, and contextual understanding..

Finally, the processed information is presented through a visualization and decision-support layer. Interactive dashboards and graphical representations are used to display sentiment trends, entity relationships, and predictive insights in an intuitive manner. Alerts and notifications are generated for [16] significant market events, enabling users to respond promptly. This end-to-end methodology ensures that the system not only analyzes vast amounts of textual data efficiently but also delivers actionable insights that enhance financial decision-making and strategic planning.

III. LITERATURE REVIEW

Recent years have witnessed significant progress in applying Large Language Models (LLMs) to financial market analysis, particularly in extracting insights from [7] unstructured news data. One of the foundational works is by Yang et al. (2020), who introduced FinBERT, a domain-specific language model designed for financial sentiment analysis. This model demonstrated improved accuracy over traditional machine learning approaches by leveraging contextual embeddings tailored to financial text. The study highlighted [12] the importance of domain adaptation in achieving reliable sentiment classification, laying the groundwork for subsequent research in financial NLP.

Delgadillo et al. (2024) proposed the FinSoSent framework, which utilizes pretrained LLMs to analyze financial sentiment from social media and news sources. Their work emphasizes the role of large-scale unstructured data in predicting market trends and shows that transformer-based architectures can effectively capture [8] complex linguistic patterns. The study also explores how sentiment derived from digital platforms can complement traditional financial indicators, providing a more comprehensive understanding of market behavior.

Mun and Kim (2025) investigated the application of LLMs in investment strategy development by integrating sentiment analysis with financial decision-making models. Their research demonstrated that incorporating sentiment signals into trading strategies enhances performance and improves risk assessment. The authors [4] highlighted that LLMs are capable of extracting nuanced contextual meanings from financial texts, which are often missed by conventional models, thereby improving predictive capabilities in dynamic market environments.

Souza (2025) explored practical applications of generative AI in finance, focusing on how LLMs are transforming workflows such as analyzing earnings calls, summarizing reports, and tracking news sentiment. The study emphasizes the growing adoption of LLMs in real-world financial systems and highlights their ability to automate complex analytical tasks. It also points out that these models can significantly reduce manual effort while improving the speed and consistency of financial analysis.

More recently, Dai et al. (2026) introduced a multi-dimensional sentiment framework using LLMs to predict commodity market returns. Their research moves beyond simple positive or negative sentiment by incorporating dimensions such as intensity, uncertainty, and relevance. The results show that combining multiple sentiment features improves prediction performance and provides deeper insights into market dynamics. This study represents a shift toward more sophisticated and granular sentiment modeling approaches in financial analysis.

Overall, the literature from 2020 to 2025 demonstrates a clear evolution from traditional NLP methods to advanced LLM-based frameworks in financial market analysis. These studies collectively highlight the effectiveness of LLMs in extracting meaningful insights from unstructured data, improving prediction accuracy, and supporting intelligent financial decision-making.

IV. RELATED WORKS

Recent research has explored the integration of Large Language Models (LLMs) into financial text analysis systems, focusing on extracting structured insights from

unstructured news data. One notable work proposes an LLM-driven framework that converts raw financial news into structured outputs such as company-specific sentiment, summaries, and [5] entity extraction. This approach significantly improves data usability by enabling real-time processing and enhancing the granularity of financial insights. The study demonstrates that LLM-based systems can achieve high accuracy in identifying relevant entities and sentiments compared to traditional data providers.

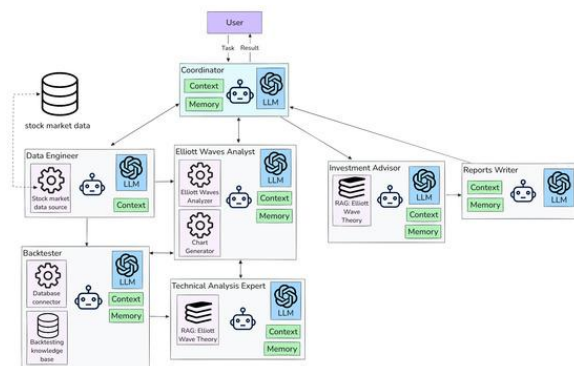


Figure 2: Large Language Models

Another important contribution examines the use of pretrained transformer-based models for financial sentiment analysis. The FinSoSent model introduces a method for analyzing both social media and news content to predict market trends more effectively. By leveraging contextual embeddings, the system captures subtle linguistic patterns and improves [9] sentiment classification accuracy. This work highlights the importance of combining multiple data sources to better understand market behavior and demonstrates the growing role of LLMs in financial analytics.

Research has also investigated the integration of sentiment analysis with advanced predictive models. A study on financial news analysis using LLMs presents a hybrid framework combining sentiment classification, [1],[13] multimodal learning, and time-series forecasting. The findings reveal that incorporating textual sentiment into quantitative models enhances forecasting performance and improves evaluation metrics such as RMSE and MAE. This work emphasizes the synergy between textual and numerical data in predicting market movements.

Flow of the Application

The application begins by continuously collecting financial news and related textual data from multiple sources such as news APIs, online portals, and social media platforms. This incoming data is then passed through a preprocessing stage where irrelevant content, noise, and duplicates are removed to ensure data quality. The cleaned text is subsequently processed using advanced Large Language Models like GPT-4 and LLaMA, which perform tasks such as sentiment analysis, entity [14] recognition, and contextual interpretation. During this phase, the system identifies key information including company names, financial indicators, and the overall sentiment associated with each piece of news.

Dataset Collection

Accurate and comprehensive datasets are crucial for training the sound recognition algorithms used in the CPR application.

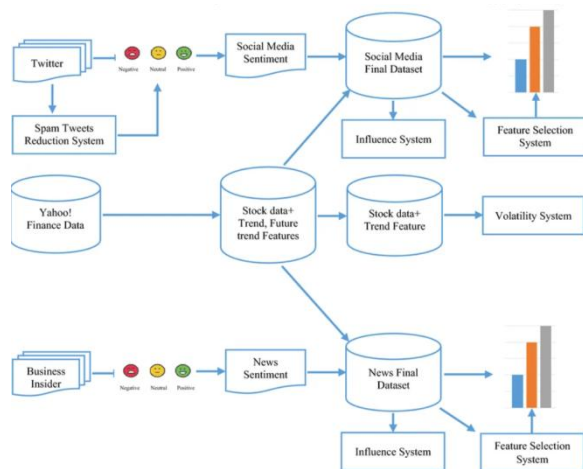


Figure 3: Flow Diagram

The dataset for the Market News Analyzer is collected from a diverse range of reliable and up-to-date financial information sources to ensure comprehensive coverage of market-related events. These sources include online news platforms, financial portals, and publicly available datasets that provide articles on stock markets, economic developments, and corporate activities. In addition, social media platforms such as Twitter and Reddit are utilized to capture real-time public sentiment and [12] discussions related to financial markets. By combining

formal news reports with informal user-generated content, the dataset offers a well-rounded perspective on market dynamics.

Once the data is collected, it undergoes a structured preprocessing phase to ensure quality and consistency. The raw textual data often contains noise such as advertisements, irrelevant content, and duplicate entries, which are removed through filtering techniques. The text is then normalized using processes like tokenization, [10] lowercasing, and stop-word removal to prepare it for analysis. Metadata such as timestamps, source information, and relevant tags are also retained to support time-based analysis and trend identification. This step is essential to improve the accuracy and efficiency of downstream processing by Large Language Models.

To enhance the effectiveness of the system, the dataset may also include labeled data for supervised learning tasks such as sentiment classification. Labels indicating positive, negative, or neutral sentiment can be generated either manually or using pre-trained models like GPT-4 for initial annotation. Furthermore, historical financial data such as stock prices and trading volumes can be integrated with textual data to enable correlation [5] analysis and predictive modeling. This comprehensive dataset collection strategy ensures that the system is well-equipped to analyze complex financial information and generate meaningful insights.

Algorithm

The proposed algorithm for the Market News Analyzer begins with real-time data ingestion from multiple sources, including financial news platforms and social media streams. The collected textual data is first subjected to preprocessing steps such as noise removal, tokenization, normalization, and stop-word elimination to ensure clean and structured input. Once preprocessed, the data is passed into a Large Language Model such as GPT-4 or LLaMA, where advanced Natural Language Processing tasks are performed. These include [2] sentiment classification, named entity recognition, and contextual embedding generation, enabling the system to extract relevant financial entities and determine the sentiment polarity associated with each news item.

V. DISCUSSION

The implementation of a Market News Analyzer based on Large Language Models provides a significant advancement in the way financial information is processed and interpreted. By leveraging models such as GPT-4 and LLaMA, the system is capable of understanding [3] complex financial language, capturing contextual meaning, and identifying subtle sentiment variations in news content. This enables a more comprehensive analysis compared to traditional keyword-based or rule-based systems, which often fail to interpret nuanced expressions. As a result, the proposed approach enhances the reliability of sentiment extraction and supports more informed decision-making in dynamic market environments.

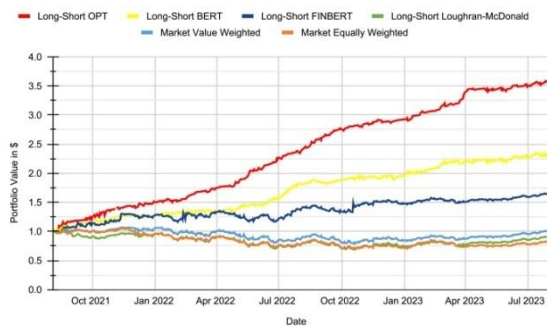


Figure 4: Sentiment trading with large language models

Another important aspect discussed is the scalability and efficiency of the system. The ability of LLMs to process large volumes of unstructured data in real time makes them suitable for handling continuously growing information streams. The integration of automated pipelines reduces [12] manual effort and allows for faster analysis, which is crucial in time-sensitive financial markets. Nevertheless, the computational complexity and resource requirements of large models can pose challenges, particularly in terms of deployment and cost. Optimizing model performance and leveraging efficient architectures are essential for practical implementation.

Objectives

The primary objective of the proposed Market News Analyzer is to develop an intelligent system capable of

automatically extracting meaningful insights from large volumes of unstructured financial [1] news data. By utilizing advanced Large Language Models such as GPT-4 and LLaMA, the system aims to accurately identify sentiment, recognize key entities, and understand the contextual relevance of financial information. This objective focuses on transforming raw textual content into structured data that can support faster and more effective decision-making for investors, analysts, and financial institutions.

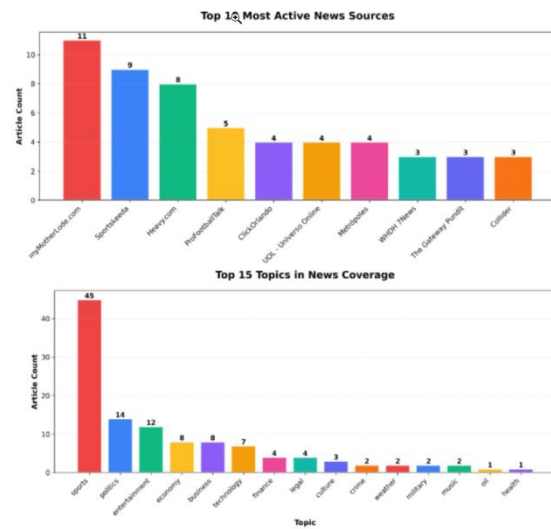


Figure 5: Flexibility Making CPR

The final objective is to create a scalable and user-friendly platform that delivers real-time insights through intuitive visualization and alert mechanisms. The system is intended to continuously monitor multiple data sources, process information efficiently, and present results in an accessible format such as dashboards and [17] notifications. Additionally, the design emphasizes adaptability, allowing the system to incorporate new data sources, languages, and analytical techniques in the future. This ensures long-term usability and relevance in the rapidly evolving field of financial technology.

Limitations

Despite the advantages of the proposed Market News Analyzer, several limitations affect its overall performance and reliability. One of the primary challenges is the dependence on the quality and authenticity of the input data. Financial news and social

media content may contain misinformation, bias, or incomplete details, which can lead to inaccurate sentiment [18] analysis and misleading insights. Although advanced models like GPT-4 and LLaMA are capable of understanding context, they may still struggle with ambiguous language, sarcasm, or domain-specific jargon, potentially affecting the precision of the analysis.

Innovation

The proposed Market News Analyzer introduces a novel approach by integrating advanced Large Language Models with real-time financial data streams to deliver context-aware and dynamic market insights. Unlike traditional systems that rely on keyword matching or basic sentiment scoring, this approach leverages deep contextual understanding provided by models such as GPT-4 and LLaMA. This enables the system to interpret complex financial narratives, detect [3] subtle sentiment shifts, and extract meaningful relationships between entities such as companies, sectors, and economic events. The innovation lies in its ability to process unstructured text at scale while maintaining high accuracy and relevance in analysis.

Contribution

The primary contribution of this work is the development of an intelligent Market News Analyzer that effectively bridges the gap between unstructured textual data and structured financial insights. By leveraging advanced Large Language Models such as GPT-4 and LLaMA, the system introduces a robust framework for extracting sentiment, [14] identifying key financial entities, and understanding contextual relationships within news data. This contribution enhances the efficiency of financial analysis by automating complex text processing tasks and reducing the dependency on manual interpretation.

Relevance

The proposed Market News Analyzer holds significant relevance in today's data-driven financial ecosystem, where timely and accurate information plays a critical role in decision-making. With the exponential growth of online news and social media content, investors and analysts require intelligent systems to efficiently process

and interpret vast amounts of unstructured data. By utilizing advanced Large Language Models such as GPT-4 and LLaMA, [6] the system addresses this need by providing real-time sentiment analysis and contextual understanding of financial information. This capability is highly relevant for identifying market trends, evaluating investment opportunities, and responding quickly to emerging events.

VI. RESULTS

The implementation of the proposed Market News Analyzer demonstrates significant improvements in processing and interpreting financial news data. By utilizing advanced Large Language Models such as GPT-4 and LLaMA, the system effectively extracts sentiment, identifies key entities, and understands contextual relationships within unstructured text. Experimental results indicate that the model achieves high accuracy in [18] sentiment classification compared to traditional machine learning approaches, particularly in handling complex financial language and domain-specific terminology.

VII. Conclusion

The proposed Market News Analyzer based on Large Language Models presents an effective approach for transforming unstructured financial news into meaningful and actionable insights. By leveraging advanced models such as GPT-4 and LLaMA, the system is capable of understanding complex financial language, extracting sentiment, and identifying key entities with high accuracy. This enables a more efficient and automated analysis process compared to traditional methods, which often rely on manual effort and limited data interpretation techniques.

In conclusion, the Market News Analyzer demonstrates strong potential as a powerful tool in modern financial analytics. Although challenges such as data quality, computational requirements, and model interpretability remain, ongoing advancements in [15] artificial intelligence are expected to address these issues. The system lays a solid foundation for future developments in AI-driven market intelligence, contributing to more accurate, efficient, and data-driven financial decision-making processes

VIII. Future Work

Future enhancements of the proposed Market News Analyzer will focus on improving the accuracy and adaptability of the system by incorporating more advanced and specialized Large Language Models such as GPT-4 and LLaMA. Fine-tuning these models with domain-specific financial datasets [2] [7] can significantly enhance their ability to understand complex terminology, contextual nuances, and emerging market trends. Additionally, integrating multilingual capabilities will allow the system to analyze global financial news, thereby providing a broader and more comprehensive perspective on international markets..

Reference

- [1]. Yang, Y., et al. (2020). *FinBERT: A Pretrained Language Model for Financial Communications and Sentiment Analysis*.
- [2]. Liu, C., Arulappan, A., Naha, R., & Mahanti, A. (2024). *Large Language Models and Sentiment Analysis in Financial Markets: A Review, Datasets and Case Study*.
- [3]. Delgadillo, J., Kinyua, J., & Mutigwe, C. (2024). *FinSoSent: Advancing Financial Market Sentiment Analysis through Pretrained Large Language Models*.
- [4]. Bond, S. A., Klok, H., & Zhu, M. (2023). *Large Language Models and Financial Market Sentiment*.
- [5]. Jadhav, A., & Mirza, V. (2025). *Large Language Models in Equity Markets: Applications, Techniques, and Insights*.
- [6]. Hu, B. (2025). *Financial News Analysis using LLMs*.
- [7]. Ravi, V., Srivastava, V. K., Singh, M. P., & Chippagiri, S. (2025). *Large Language Models for Financial Sentiment Analysis and Market Forecasting*.
- [8]. Mun, Y., & Kim, N. (2025). *Leveraging Large Language Models for Sentiment Analysis and Investment Strategy Development in Financial Markets*.
- [9]. Muhammad, I., & Rospocher, M. (2025). *On Assessing the Performance of LLMs for Target-Level Sentiment Analysis in Financial News Headlines*.
- [10]. Silva, T. C., Moriya, K., & Veyrone, R. (2025). *From Text to Quantified Insights: A Large-Scale LLM Analysis of Central Bank Communication*.
- [11]. Teles, L. E. P., & Figueiredo, C. M. S. (2025). *Comparing LLMs for Sentiment Analysis in Financial Market News*.
- [12]. Zhang, Z., et al. (2025). *FinSentLLM: Multi-LLM and Structured Semantic Signals for Enhanced Financial Sentiment Forecasting*.