iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VIII Month of publication: August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73799

Automated Market Intelligence: A Multi-Agent Approach with MCP

Amaan Mithani[1], Amaan Ansari[2], Gaurav Ghop[3]
Department of Electrical and Computer Engineering, New York University

Abstract: The sheer volume and velocity of information in financial markets create significant challenges for timely and accurate analysis. This paper presents a multi-agent system that uses relation extraction to derive actionable intelligence from financial news, corporate press releases, and market filings. The proposed Agentic AI system combines four agents: (i) a machine learning agent for sentiment analysis (logistic regression), (ii) a corporate profile agent for baseline fact-checking (which relies on named entity recognition), (iii) a narrative consistency agent (using LLM prompt engineering), and (iv) a real-time market data analyzer that extracts relational triplets for claim verification.

The system is orchestrated via the Model Context Protocol (MCP), offering shared context and live learning across components. Results demonstrate that the multi-agent ensemble achieves 95.3% accuracy with an F1 score of 0.964 in identifying verifiable market claims, significantly outperforming individual agents and traditional approaches. The weighted aggregation method, mathematically derived from individual agent misclassification rates, proves superior to algorithmic threshold optimization. The modular architecture makes the system easily scalable while maintaining details of the analytical processes.

Keywords: Market Intelligence, Agentic AI, Model Context Protocol (MCP), Relation Extraction (RE), Financial NLP, LLM

I. INTRODUCTION

The manipulation of information has always been a powerful tool for influencing market perceptions and securing a competitive advantage. A key challenge in modern market analysis is distinguishing genuine signals from noise in the torrent of financial data. Throughout history, speculative information has been used by companies to inflate valuations, manage reputations, or generate investment interest. Flashy claims about new products or partnerships can spread quickly, impacting trading decisions before the underlying facts are fully vetted.

In the study of financial information, data is generally split into categories: verifiable facts (e.g., reported earnings), forward-looking statements (projections with inherent uncertainty), and corporate narratives (biased information used to promote a strategic agenda). While these categories differ in intent and certainty, from a computational perspective, they must all be considered when evaluating a company's announcements. They all obey the same constraints, being bound to language patterns, context, and source credibility.

Speculative or misleading financial information is often based on fabricating new relationships between entities (e.g., companies, products, technologies) to excite investors. Relation Extraction (RE) identifies and classifies semantic relationships between entities mentioned in text. We aim to enhance market intelligence by applying RE to financial news, focusing on headlines and executive summaries where key claims are often concentrated. These texts are crucial because they are the primary tools for shaping investor opinion and market sentiment.

II. RELATED WORK

Our work builds upon key advancements in financial NLP, from foundational sentiment analysis to modern agentic systems. Foundational research, such as the influential financial sentiment dictionaries developed by Loughran and McDonald (2011), established the need for domain-specific tools. This field has since evolved to include more complex methods like relation extraction (RE), where researchers use deep learning to construct financial knowledge graphs from news events, as demonstrated by Ding et al. (2019). The recent advent of Large Language Models (LLMs) has marked a paradigm shift. This is exemplified by the development of massive, domain-specific models like BloombergGPT (Wu et al., 2023) and open-source alternatives like FinGPT (Yang et al., 2023). These monolithic models excel at a range of financial tasks but can be opaque. Our approach diverges by creating a cooperative system of smaller, specialized agents, drawing inspiration from the potential of "generative agents" to simulate complex behaviors (Park et al., 2023).

While prior works focus on individual tasks like sentiment or single-model performance, our primary contribution is the synthesis and orchestration of these varied techniques. We integrate them into a unified, multi-agent framework that performs robust, evidence-based claim verification.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

III.USED DATASETS

The system is trained on two main types of datasets: financial news archives (from sources like Reuters and Bloomberg) and SEC EDGAR filings. These datasets complement one another; the news provides real-time market sentiment and narratives, while the filings offer structured, regulated corporate data. For source credibility, the system uses a curated dataset ranking financial news outlets based on their historical accuracy and market impact, analogous to the Iffy News project for general news. The training data is a balanced set of verifiable claims (e.g., product launches confirmed by multiple sources) and speculative or unverifiable statements, representing roughly 10% of the initial data. Table. 1 details the composition of our balanced dataset, breaking down the number of "Verifiable" and "Speculative" claims across the training, validation, and test splits.

Table. 1: Class Distribution Across Training, Testing and Validation Dataset

Class	Training Set	Validation Set	Test Set	Total
Verifiable Claim	2,100	450	500	3,050
Speculative / Unverifiable Claim	2,100	450	500	3,050
Total Samples	4,200	900	1,000	6,100

IV. AGENTIC MARKET INTELLIGENCE ENGINE

The solution has three parts: (1) the market intelligence engine based on Agentic AI; (2) a source and report recommendation system; and (3) a search engine query generator for deeper investigation. The second and third modules allow for a human-in-the-loop, enabling analysts to direct the investigation. The system is composed of four agents, an orchestrator, and an aggregator that combines their outputs into a single analysis as shown in Fig. 1.

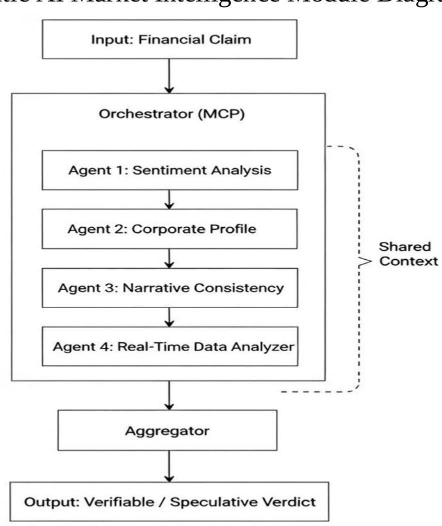
A. Sentiment Analysis Agent

This agent provides a fast prediction of market sentiment surrounding a piece of text. It is based on a Hashing Vectorizer and a Stochastic Gradient Descent (SGD) classifier trained on financial texts labeled as positive, negative, or neutral. The logistic regression model outputs probabilities that are used to classify the input's sentiment and calculate the agent's confidence score. This agent excels when analyzing news from sectors it was trained on but can misinterpret nuanced language in unfamiliar domains.

B. Corporate Profile Agent

The main idea is that a veracity benchmark can be established by checking claims against established, widely accepted corporate data. The agent uses Named Entity Recognition (NER) to extract company names, products, and executives from an input text. It then queries trusted knowledge bases (e.g., Wikipedia, financial data providers) to retrieve summaries. A claim's credibility is assessed based on the keyword overlap between the input text and the retrieved profiles, assuming factual statements align more closely with established records. For example, a claim about a new CEO would be checked against the company's official leadership roster.

Fig. 1 Agentic AI Market Intelligence Module Diagram



C. Narrative Consistency Agent

Singh et al. observed that misleading articles often exhibit lower textual coherence. In a market context, a company's announcements should be logically consistent with its past statements and overall strategy. A sudden, unexplained pivot can be a red flag. This agent uses a large language model (LLM) to grade the narrative coherence of a new press release or announcement against the company's previous quarterly reports and investor calls. It uses the "Act As" prompt pattern to instruct the LLM to behave as a financial analyst looking for strategic inconsistencies.

D. Real-Time Market Data Analyzer Agent

Effective analysis models rely heavily on large volumes of real-time data to identify emerging trends accurately. This agent uses a triplet extraction system on data fetched from the web in real-time. It queries the web for content related to a claim, uses a relation extraction model to pull out structured triplets (e.g., Company X, acquires, Company Y), and compares these web-derived triplets to those from the original input text[4]. The Llama3 8B model is then prompted to evaluate the evidence and provide a structured output with a label (e.g., "Verified," "Conflicting," "Speculative"), a confidence level, and an explanation. This agent has the best single-agent performance due to its real-time access to market information. Table. 2 shows the performance of each individual agent compared to the final, aggregated output of the multi-agent ensemble. The ensemble's superior F1-Score highlights the benefit of combining the agents' complementary strengths.

Table. 2: Performance Results of the Multi-Agent System

Agent	Accuracy	Precision	Recall	F1-Score
Sentiment Analysis Agent	78.5%	0.81	0.77	0.79
Corporate Profile Agent	84.1%	0.85	0.83	0.84
Narrative Consistency Agent	76.2%	0.79	0.75	0.77
Real-Time Market Data Analyzer	92.8%	0.94	0.92	0.93
Multi-Agent Ensemble (Final System)	95.3%	0.96	0.97	0.964

V. AGENT ORCHESTRATION AND HUMAN-IN-THE-LOOP

A. Agent Orchestration and Aggregation

The system's multi-agent architecture requires careful orchestration to ensure context is shared and leveraged effectively. The orchestrator, using the Model Context Protocol (MCP), coordinates the agents in a sequence, allowing each to update a shared context that is passed to the next agent in the pipeline. The Aggregator then calculates a weighted average of each agent's output to produce a final verdict. The weights are determined by each agent's historical misclassification rate on market data, with the final weights being:

Sentiment Analysis Agent (19%), Corporate Profile Agent (23%), Narrative Consistency Agent (17%), and Real-Time Market Data Analyzer (41%).

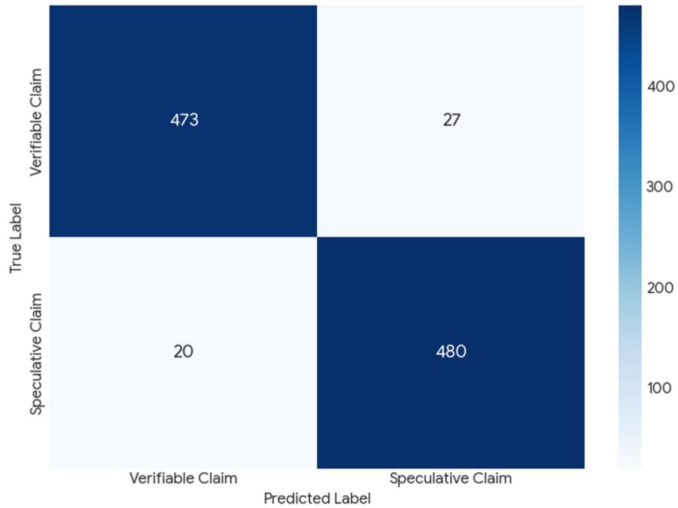
The weights reflect each agent's role. The Real-time Market Data Analyzer has the highest weight because it accesses the most current and comprehensive data, which is paramount in finance. The Narrative Consistency Agent has the lowest relevance, as strategic shifts, while important, are less definitive than hard data.

B. Human-In-The-Loop Sub-system

To empower human analysts, the system includes a two-track recommendation module: one for ranking relevant web-based sources and another for retrieving official filings and academic literature. It scores web sources using a combination of semantic similarity to the user's query and the source's predefined trustworthiness score. It also queries academic databases like arXiv and financial databases like EDGAR to find supporting or refuting evidence, allowing an analyst to quickly access primary source documents for deeper investigation.

Fig. 2 demonstrates a comprehensive performance comparison by plotting four key metrics for each agent and the final ensemble. It clearly demonstrates that the integrated multi-agent system is superior, outperforming every individual component across all measures of accuracy, precision, recall, and F1-score.

Fig. 2 Confusion Matrix for Multi-Agent Ensemble



VI. DISCUSSION

The proposed Agentic AI system achieves 95.3% accuracy and an F1 score of 0.964 in identifying verifiable market claims. However, several practical challenges remain. The system's dependence on web-sourced data creates both opportunities and risks. While real-time API access helps detect emerging market trends, its effectiveness drops if the APIs fail or if search results are influenced by coordinated market manipulation (e.g., "pump and dump" schemes).

Performance across different market sectors (e.g., tech vs. healthcare) remains an important area for investigation. Political news might influence tech stocks differently than pharmaceutical trial results affect biotech stocks, presenting unique analytical challenges. The current architecture has not been tested against adversarial attacks, where malicious actors might try to exploit the system's patterns to promote a false narrative.

Despite these limitations, the multi-agent approach offers clear advantages for market intelligence. Its modular design allows for continuous improvement of individual components. The use of diverse methods, from sentiment analysis to real-time data verification, provides robust cross-validation against different types of market information. Finally, having multiple agents explain their reasoning improves transparency, helping human analysts make more informed and confident decisions.

VII. CONCLUSION

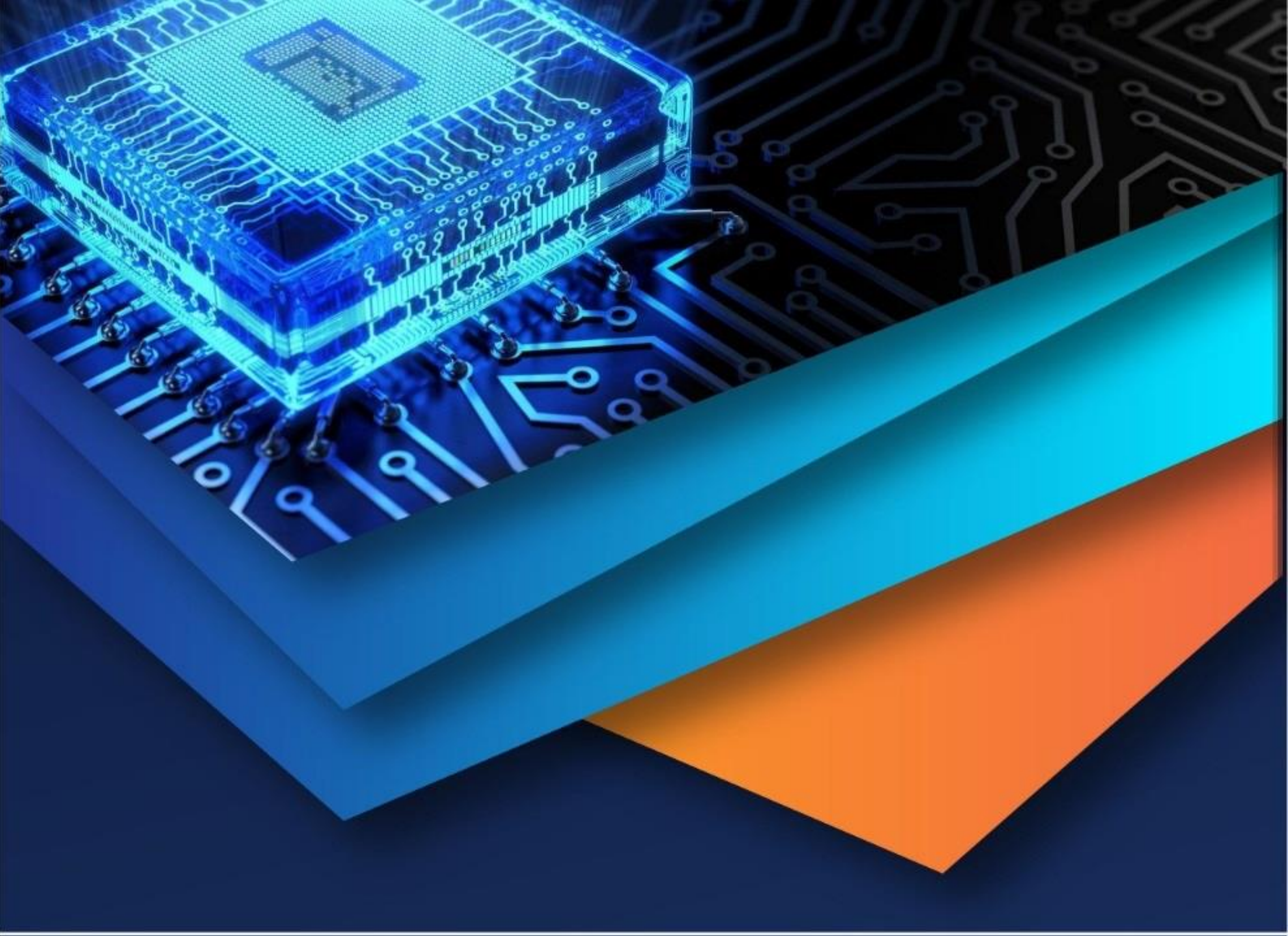
Deriving actionable market intelligence through relation extraction is a highly complex task, posing not just technical but also semantic and contextual problems. Agentic methods have proven to perform better, as multiple specialized agents work together like a team of financial analysts, each bringing a different expertise to bear on the problem.

A key takeaway is that complementarity is invaluable in complex analytical tasks. No single technique, whether a lightweight sentiment classifier or a massive LLM, can handle the task of market intelligence alone. In line with the MCP, we learn that context is a "first-class citizen". Incremental learning and shared context passing allowed each agent to build upon the findings of its predecessors, creating a richer, more nuanced final analysis. Our work demonstrates that combining diverse AI approaches through MCP orchestration can achieve significant improvements in analytical accuracy for market intelligence. As market dynamics grow more sophisticated, the modularity and human-in-the-loop features of the proposed solution position it well for future enhancements and real-world deployment in financial institution

REFERENCES

[1] Ding, L., Liu, Y., Zhang, Z., & Li, B. (2019). "Extracting Structured Event Knowledge from Financial News." EMNLP-IJCNLP.
[2] Loughran, T., & McDonald, B. (2011). "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." The Journal of Finance.
[3] Park, J. S., et al. (2023). "Generative Agents: Interactive Simulacra of Human Behavior." UIST.
[4] Wu, S., et al. (2023). "BloombergGPT: A Large Language Model for Finance." arXiv preprint.
[5] Yang, H., Liu, X. Y., & Wang, C. (2023). "FinGPT: Open-Source Financial Large Language Models." arXiv preprint.
[6] Hou, T. E. (2023). Model context protocol. The MCP Foundation.
[7] Li, G., Madaan, A., Zettlemoyer, L., & Yih, W. T. (2023). CAMEL: Communicative agents for "mind" exploration of large language model society. In Advances in Neural Information Processing Systems (NeurIPS).

[8] Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST).
[9] Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. The Knowledge Engineering Review, 10(2), 115–152.
[10] Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In Proceedings of the 11th International Conference on Learning Representations (ICLR).



10.22214/IJRASET

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY