

Dear Editors and Guest Editors,

I am pleased to submit my manuscript "Predicting Market Reactions to News: An LLM-Based Approach Using Spanish Business Articles" for consideration in the Special Issue on Artificial Intelligence and Finance of the *International Review of Financial Analysis*. This paper introduces a novel methodology that leverages Large Language Models (LLMs) to analyze how markets process and react to business news. In particular, we develop a structured schema that guides LLMs to systematically identify and classify firm-specific economic shocks in news articles according to their type, magnitude, and direction. By translating these economically-motivated classifications into trading signals, we are able to construct a trading strategy that generates consistent profits out-of-sample, demonstrating that our LLM-based methodology can effectively identify market trends while maintaining interpretability and economic relevance.

This paper directly aligns with the scope of the special issue, as it applies state-of-the-art generative AI methodologies to asset pricing and market efficiency. By integrating economic frameworks into LLM-guided analysis, this work advances the practical application of generative AI to identify persistent market patterns, offering robust and interpretable insights.

#### **Key Contributions of the Paper:**

1. *Methodological Innovation:* We propose a new framework that moves beyond traditional text analysis methodologies by embedding economic structure into LLM-based text classification. This approach enhances both the interpretability and the economic relevance of the extracted signals.
2. *Predictive Superiority and Profitability:* Our methodology demonstrates superior predictive power relative to conventional vector embedding techniques. These gains translate into consistent out-of-sample profitability, supported by transparent, interpretable, and durable trading signals.
3. *Revealing Persistent Patterns in Market Reactions:* Guided by economic frameworks, LLMs can effectively identify persistent patterns in how markets react to diverse types of firm-specific news, thereby strengthening our understanding of market efficiency and the informational content of news events.

#### **Incorporation of Feedback from Recent Conferences:**

This research has been presented at several academic conferences where it benefited from candid discussions and valuable feedback that enhanced both its methodological rigor and practical relevance. Most recently, it benefited from presentations and feedback at the Generative AI in Finance Conference held in Montréal (October 2024), the Banking & Finance seminar at CEMFI (October 2024) and the 3rd Contemporary Issues in Financial Markets and Banking (January 2025). We have incorporated all key feedback received:

- **Expanded Risk Metrics:** We have now included Sortino Ratio, Maximum Drawdown, Calmar Ratio, Skewness, Excess Kurtosis, Value at Risk (VaR), and Conditional VaR (CVaR) to provide a comprehensive evaluation of trading strategy performance (see Table 5, Table A.4).

- Enhanced Methodology Interpretation: We now provide a clearer interpretation of the embedding-based methodology to improve comprehension of the technical underpinnings (Section 4.1.2).
- Trading Intensity and Transaction Costs: We analyze trading intensity differences between KMeans and LLM methodologies, incorporating conservative estimates of transaction costs (Section 5.3, Appendix Section A.8).
- Foundational Explanation of LLMs: We introduce and contextualize LLMs and their evolution prior to detailing our LLM-based approach (Sections 4.2.1 and 4.2.2).
- Illustrative Examples: We include example articles to showcase how the methodology operates in practice (Examples 1 and 2).
- Benchmarking Rationale: We rigorously justify our choice of vector embeddings as the benchmark and explain why sentiment analysis or topic modeling approaches are less suitable comparators (Section 4.1.1, Appendix Section A.7).

**Data and Transparency:**

The dataset, sourced from Dow Jones Newswires, is proprietary and cannot be publicly shared. Interested readers may access the same data from Dow Jones Newswires for a fee. In the manuscript we ensure that the methodology and findings are rigorously documented and fully reproducible.

The manuscript has not been published or submitted elsewhere. We have no conflicts of interest to declare. The research has been conducted with financial support from Banco de España, as acknowledged in the manuscript. We propose no exclusions for reviewers.

Given the paper's direct alignment with the special issue's focus and its comprehensive integration of economic structure, LLMs, and robust empirical analyses we believe it will make a valuable contribution to this special issue. We greatly appreciate your time and consideration and look forward to your response. Should you require any additional information, please do not hesitate to contact me.

Sincerely,

*Jesús Villota*