

Reviewer 2 Comments - Round 2

Manuscript: NarrativeBreak: Integrating Structural Break Detection with Multi-Source NLP Signals for Dynamic Portfolio Optimization **Recommendation:** Minor Revision **Confidence:** High

Summary

The revision substantially improves the paper. The authors have been responsive to all major concerns, and I appreciate the honest treatment of statistical significance after multiple testing correction. The real-data case study in Appendix B, while limited in scope, provides valuable evidence that the framework can work with actual text data.

I have a few remaining comments, mostly about presentation and framing.

Minor Issues

1. Abstract Framing

The abstract still emphasizes the raw p-value of 0.034 without mentioning the Bonferroni-adjusted p-value of 0.17. Given the importance of proper statistical inference, the abstract should be updated to reflect the corrected significance level or use more cautious language.

Suggested revision: Change “statistically significant outperformance” to “suggestive evidence of outperformance (Bonferroni-adjusted $p = 0.17$)” or simply describe the magnitude of improvement without claiming significance.

2. Conclusion Section

Similarly, Section 7 (Conclusion) should include a brief note about the marginal significance. The current framing could be interpreted as overclaiming.

3. GPT-4 Discussion

The discussion of GPT-4 in Section 5.4 is helpful but could be expanded slightly. Given the rapid pace of LLM development, please add: - The specific GPT-4 version tested (e.g., gpt-4-0314) - Note that costs have decreased substantially since testing (if applicable) - Acknowledgment that open-source alternatives (Llama, Mistral) may offer better cost-accuracy tradeoffs

4. Real-Data Case Study Limitations

Appendix B is valuable but the limitations should be stated more explicitly: - GDELT data quality varies by source - Entity recognition was performed manually for this case study - The 2-month window is too short for regime detection to work optimally

5. Minor Typos

- Page 7, line 23: “behaviour” should be “behavior” (assuming US English style)

- Page 12, Table 5: Missing period after “days”
 - Reference [24]: Journal name appears abbreviated inconsistently
-

Assessment

The paper is now close to publication quality. The authors have demonstrated appropriate scholarly rigor by acknowledging limitations and providing honest statistical reporting. The contribution—showing that narrative signals can provide early warning of regime changes—is meaningful and timely.

I recommend **Minor Revision** to address the framing issues noted above. These are straightforward changes that should not require re-review.

Note: This is a simulated review for demonstration purposes.