Enhancing Financial Sentiment Analysis through Multimodal Fusion for Retrieval Augmented Generation

Sagnik Nandi
Georgia Institute of Technology
snandi41@gatech.edu

Chittebbayi Penugonda
Georgia Institute of Technology
cpenugonda3@gatech.edu

Saketh Budideti
Georgia Institute of Technology
sbudideti3@gatech.edu

Joshua Zhang
Georgia Institute of Technology
jzhang3346@gatech.edu

Veer Guda
Georgia Institute of Technology
vguda6@gatech.edu

Arnav Tripathi
Georgia Institute of Technology
atripathi83@gatech.edu

I. Problem Definition and Motivation

Modern approaches in sentiment analysis rely on Large Language Models (LLMs) to assess textual data and categorize it into sentiment classes (e.g., positive, negative, neutral) [6]. However, most online sentiment often appears more exaggerated than the underlying intent of the author, influenced by heightened social incentives to express polarized views [source needed]. Within financial contexts, both general-purpose language models (such as GPT 4.0, BERT, LLaMA-7B) and domain-specific models (like FinBERT, BloombergGPT, FinGPT) often fail to possess sufficient context to moderate the fluctuating nature of online sentiment for precise analysis and categorization.

In response to this challenge, researchers have devised various techniques utilizing Retrieval Augmented Generation (RAG) systems and datasets. These approaches allow models to access contextual repositories, empowering LLMs with the latest data while circumventing extensive training processes. Nonetheless, despite the efficacy of modern RAG implementations in integrating up-to-date information, they might neglect several critical microeconomic facets concerning individual companies, necessitating further refinement and specificity.

II. State of the Art and Limitations

When looking for a state of the art, there currently does not exist any dataset which measures and metrics in the neighborhood of our dataset. The closest apparent benchmark appears to be *TAT-QA: A Question Answering Benchmark on a Hybrid of Tabular and Textual Content in Finance*, currently recognized as the most robust tabular and textual dataset for analysis [cite]. Within this benchmark, researchers seamlessly combine semi-structured tables featuring a range of financial data alongside at least two accompanying paragraphs. These paragraphs provide both descriptive and analytical insights into the table's contents, with the aim of enriching the model's contextual comprehension and strengthening the overall accuracy of language model responses.

III. Dataset Direction

In order to adapt the dataset used in the TAT-QA paper for sentiment analysis, we propose a similar tabular and textual data pair combined with textual sentiment from various online media sources (Twitter, Seeking Alpha, Yahoo Finance, etc.). The labelled portion of our dataset is the adjusted sentiment score of the online-sourced textual sentiment on a scale of 1 to 5, corresponding to very bearish sentiment, bearish sentiment, neutral sentiment, bullish sentiment, very bullish sentiment, respectively.

IV. References

[1] Gartner Glossary, "Definition of sentiment analysis - finance glossary - gartner," 2023

[2] Dogu Araci, "FinBERT: Financial sentiment analysis with pretrained language models," in arXiv preprint arXiv:1908.10063, 2019

[3] Sahar Sohangir, Dingding Wang, Anna Pomeranets, and Taghi M Khoshgoftaar, "Big data: Deep learning for financial sentiment analysis," Journal of Big Data, vol. 5, no. 1, pp. 1–25, 2018

[4] Hongyang Yang, Xiao-Yang Liu, and Christina Dan Wang, "Fingpt: Open-source financial large language models," arXiv preprint arXiv:2306.06031, 2023

[5] Zhang, Duzhen, Yahan Yu, Chenxing Li, Jiahua Dong, Dan Su, Chenhui Chu, and Dong

Yu. ”MM-LLMs: Recent Advances in MultiModal Large Language Models.” 2024. arXiv preprint, arXiv:2401.13601.

[6] Zhang, Boyu, Hongyang Yang, Tianyu Zhou, Ali Babar, and Xiao-Yang Liu. ”Enhancing Financial Sentiment Analysis via Retrieval Augmented Large Language Models.” 2023. arXiv preprint, arXiv:2310.04027.

[7] Thulke, David, Nico Daheim, Christian Dugast, and Hermann Ney. ”Efficient Retrieval Augmented Generation from Unstructured Knowledge for Task-Oriented Dialog.” 2021. arXiv preprint, arXiv:2102.04643.

[8] Zheng, L., Li, N., Chen, X., Gan, Q., & Zhang, W. (2023). Dense representation learning and retrieval for tabular data prediction. Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. https://doi.org/10.1145/3580305.3599305

[9] Xue, Zihui, and Radu Marculescu. ”Dynamic Multimodal Fusion.” 2023. arXiv preprint, arXiv:2204.00102.