Annotated Bibliography

Anany Sachan

September 25, 2025

References

[1] Dogu Araci. FinBERT: Financial Sentiment Analysis with Pre-trained Language Models, 2019.

[2] Ringki Das and Thoudam Doren Singh. Multimodal Sentiment Analysis: A Survey of Methods, Trends, and Challenges. *ACM Comput. Surv.*, 55(13s), July 2023.

Das and Singh (2023) present a comprehensive survey on the evolution of sentiment analysis from text-based methods to multimodal approaches that integrate audio, visual, and textual information. The article reviews traditional techniques such as lexicon-based and machine learning models, and highlights the role of deep learning and transformer-based architectures in advancing multimodal sentiment detection as well as identifying trends such as the growing reliance on pretrained language models and cross-modal embeddings. It discusses applications in domains like politics, healthcare, finance, and social media, while also emphasizing challenges such as sarcasm detection, inter- and intra-modality dynamics, low-resource languages, and high computational costs. The authors catalog widely used sentiment lexicons, benchmark datasets (e.g., IMDB, SST, CMU-MOSEI), and evaluation metrics, offering guidance for new researchers in the field. Published in ACM Computing Surveys and peer-reviewed, the work is highly credible and valuable for understanding both the state of the art and open research directions in multimodal sentiment analysis.

[3] Xiang Deng, Vasilisa Bashlovkina, Feng Han, Simon Baumgartner, and Michael Bendersky. LLMs to the Moon? Reddit Market Sentiment Analysis with Large Language Models. In *Companion Proceedings of the ACM Web Conference 2023*, WWW '23 Companion, pages 1014–1019, New York, NY, USA, 2023. Association for Computing Machinery.

Deng et al. (2023) investigate financial sentiment analysis in the context of Reddit, specifically the r/WallStreetBets forum. They address the lack of high-quality labeled data by employing a semi-supervised learning approach in which a large language model (LLM) generates weak sentiment labels.

These are then distilled into a smaller model optimized for deployment. The authors introduce techniques such as chain-of-thought prompting and multiple reasoning paths to improve label consistency and accuracy. Their evaluation shows that the distilled student model performs on par with supervised baselines despite requiring minimal human annotation. The paper highlights both the potential and considerations of applying LLMs in financial domains, particularly when dealing with social media language that blends financial jargon, slang, and informal discourse. Published in the peer-reviewed *ACM Web Conference Companion Proceedings*, the work is credible and timely, reflecting current advances in large language models and semi-supervised training. Its contribution lies in demonstrating a practical pipeline for leveraging LLMs to overcome data scarcity while maintaining model efficiency, offering insights into both methodological innovation and the challenges of applying NLP to complex, real-world financial discussions.

[4] Antonio Desiderio, Luca Maria Aiello, Giulio Cimini, and Laura Alessandretti. The dynamics of the Reddit collective action leading to the GameStop short squeeze. *npj Complexity*, 2(1), February 2025.

[5] Tingsong Jiang and Qingyun Zeng. Financial sentiment analysis using FinBERT with application in predicting stock movement, 2025.

[6] Saif M. Mohammad. Sentiment Analysis: Automatically Detecting Valence, Emotions, and Other Affectual States from Text. In Herb Meiselman, editor, *Emotion Measurement (Second Edition)*. Elsevier, 2021.

Mohammad (2021) provides a comprehensive survey of sentiment analysis, tracing its evolution from early polarity detection in product reviews to contemporary models that capture valence, discrete emotions, and complex affectual states. The chapter outlines major challenges—including subtle and figurative language, cross-cultural variation, and the difficulty of obtaining annotated data—while also surveying advances in machine learning such as deep neural networks, transfer learning, and contextual embeddings. It reviews resources like emotion lexicons, annotated corpora, and shared evaluation tasks, emphasizing their role in advancing the field. Importantly, Mohammad highlights ethical concerns, noting that sentiment systems can reproduce and amplify human biases, raising fairness issues in applications such as hiring, lending, and public policy. Published in Emotion Measurement 2021, the work is both highly credible and well-grounded in both theory and practice. Its breadth makes it highly relevant to projects exploring sentiment classification, as it not only synthesizes two decades of scholarship but also identifies open research problems, such as modeling figurative language and ensuring fairness in NLP systems.

[7] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? Sentiment Classification using Machine Learning Techniques. In *Proceedings of the ACL-02 Conference*

on Empirical Methods in Natural Language Processing - Volume 10, EMNLP '02, pages 79–86, USA, 2002. Association for Computational Linguistics.

> Pang, Lee, and Vaithyanathan (2002) investigate sentiment classification, aiming to categorize documents by overall opinion rather than topic. Using a large dataset of movie reviews, they compare Naive Bayes, maximum entropy, and support vector machines, finding that these algorithms significantly outperform human-selected word lists but still achieve lower accuracy than in topic-based classification. The paper highlights challenges unique to sentiment analysis, including subtle language cues, rhetorical contrasts, and the need for contextual understanding (e.g., negation). Published in the *Proceedings of EMNLP* and widely cited, the work is highly credible and foundational to sentiment analysis research. While the paper's models may seem simple by today's standards, they illustrate enduring concepts such as feature representation (bag-of-words, unigrams vs. bigrams) and classification tradeoffs. As a cornerstone in the field, this source highlights the early challenges of sentiment analysis and sets the stage for evaluating modern neural and transformer-based approaches.

[8] Jiayi Tang. Leveraging Transformer Neural Networks for Enhanced Sentiment Analysis on Online Platform Comments. In *Proceedings of the 2023 4th International Conference on Machine Learning and Computer Application*, ICMLCA '23, pages 256–260, New York, NY, USA, 2024. Association for Computing Machinery.

> Tang (2024) evaluates transformer-based neural networks for sentiment classification in large-scale online user comments. The study utilizes the Sentiment140 dataset, consisting of 1.6 million labeled tweets, to compare transformer models with earlier approaches such as support vector machines (SVMs) and long short-term memory (LSTM) networks. The results indicate that transformers significantly outperform these baselines in terms of classification accuracy and robustness, particularly in handling noisy and informal language common in social media data. The paper situates these findings within the broader development of sentiment analysis, noting the progression from rule-based and classical machine learning models to deep contextual representations. By emphasizing the advantages of transformers in capturing long-range dependencies and contextual relationships, the study highlights why they have become the dominant paradigm in NLP tasks. Presented at the peer-reviewed *International Conference on Machine Learning and Computer Application*, the work is credible and provides clear empirical evidence for the effectiveness of transformers. Its contribution lies in reinforcing the superiority of transformer-based models in sentiment analysis while situating them in the historical trajectory of methodological improvements in the field.

[9] Frank Z. Xing, Erik Cambria, and Roy E. Welsch. Intelligent Asset Allocation via Market Sentiment Views. *Comp. Intell. Mag.*, 13(4):25–34, November 2018.

Xing, Cambria, and Welsch (2018) examine how sentiment information can be incorporated into asset allocation, an area that has received comparatively little attention in finance research. They propose a hybrid neural network framework that combines evolving clustering techniques with long short-term memory (LSTM) models to derive sentiment-informed market views from social media data. These views are integrated into a Bayesian portfolio optimization framework, extending the principles of modern portfolio theory with behavioral signals. Experimental results show that portfolios enhanced with sentiment views achieve improved stability and profitability compared to traditional benchmarks. The study emphasizes not only the potential financial benefits of incorporating sentiment but also the methodological advantages of hybrid neural systems in capturing dynamic market signals. Published in the peer-reviewed *IEEE Computational Intelligence Magazine*, the article is both credible and influential in bridging computational intelligence with financial modeling. Its value lies in demonstrating how sentiment indices, often studied in isolation, can be operationalized into investment decision-making. The paper highlights a methodological precedent for translating natural language signals into quantitative financial strategies, underscoring the interdisciplinary importance of sentiment analysis across computer science and finance.

[10] Lei Zhang, Shuai Wang, and Bing Liu. Deep learning for sentiment analysis: A survey. *WIREs Data Mining and Knowledge Discovery*, 8(4):e1253, 2018.

Zhang, Wang, and Liu (2018) provide a detailed survey of how deep learning has been applied to sentiment analysis, synthesizing developments up to 2018. The paper reviews key neural architectures including convolutional neural networks (CNNs), recurrent neural networks (RNNs) with gated mechanisms such as LSTM and GRU, and early applications of attention-based models. It discusses how these methods surpass traditional classifiers like SVMs and Naïve Bayes by capturing complex semantic and syntactic features. The authors also explore issues related to feature representation, highlighting the role of word embeddings, hierarchical document modeling, and hybrid systems. In addition, they identify challenges that persist in deep learning approaches, such as the heavy need for labeled data, limited ability to handle sarcasm and irony, and difficulties in cross-domain generalization. Published in the peer-reviewed *WIREs Data Mining and Knowledge Discovery*, the article is highly credible and written by leading researchers in sentiment analysis. Its contribution lies in mapping the progression from traditional models to deep neural architectures and outlining limitations that spurred subsequent innovations. The work remains an important reference for understanding how deep learning laid the groundwork for the adoption of transformer-based architectures in natural language processing.