Hypothesizing Multimodal Influence: Assessing the Impact of Textual and Non-Textual Data on Financial Instrument Pricing Using NLP and Generative AI

Karolina Bolesta
Warsaw School of Economics
Warsaw, Poland
kboles@sgh.waw.pl

Gabin Taibi
Bern University of Applied Sciences
Bern, Switzerland
gabin.taibi@bfh.ch

Codruta Mare
Babes-Bolyai University
Cluj Napoca, Romania
codruta.mare@econ.ubbcluj.ro

Branka Hadji Misheva
Bern University of Applied Sciences
Bern, Switzerland
branka.hadjimisheva@bfh.ch

Christian Hopp
Bern University of Applied Sciences
Bern, Switzerland
christian.hopp@bfh.ch

Joerg Osterrieder
Bern University of Applied Sciences
Bern, Switzerland
joerg.osterrieder@bfh.ch

Abstract

This paper presents an advanced conceptual framework for the analysis of textual data in the context of financial securities, hypothesizing that a comprehensive evaluation of events within the broader economic environment, particularly through their descriptions, significantly influences the pricing of financial instruments. This research extends beyond the traditional scope of Natural Language Processing by proposing the inclusion of non-textual data forms such as images, videos, and audio in the analysis. Further, it acknowledges the recent developments in Generative Artificial Intelligence, suggesting its application to expand the breadth of textual analysis through the generation of varied textual datasets. The hypothesis posits that the systematic analysis of these diverse multimodal textual inputs, surpassing the conventional verbal text, could enhance the decision-making process in financial asset management. This study aims to elucidate the potential effects of this methodological advancement on financial market fluctuations and outlines the most pertinent NLP methodologies for the empirical investigation of the hypothesis in future scholarly work.

1   Introduction

The exponential increase in data volume in the digital era represents a significant paradigm shift, particularly in the domain of financial market analysis. The advent of the Internet and social media platforms has catalyzed an unprecedented growth in global data volumes. From a relatively modest scale of a few dozen exabytes in 2003 Cambria and White (2014), the year marking the emergence of social media, this figure is projected to soar to approximately 180 zettabytes by 2025 Statista (2023). This rapid expansion of data presents both formidable challenges and unique opportunities for contemporary financial analysis methodologies.

Within this context, Natural Language Processing (NLP), a core facet of Artificial Intelligence, has emerged as an instrumental technology in understanding the vast and complex datasets prevalent in today's digital landscape. Recognized for its efficacy in processing extensive textual information, NLP's applications have gained considerable traction across multiple sectors, notably within the financial industry. However, the hypothesis posed in this paper extends the analytical reach beyond the consideration of traditional text-centric data. It argues for an integrative framework that combines the analysis of textual data with non-textual elements such as images, videos, and audio. This multimodal approach is essential to have a more holistic and nuanced understanding of financial market dynamics.

Furthermore, the advent and integration of Generative Artificial Intelligence (AI) in data analysis signifies a pivotal advancement in this field. The capabilities of Generative AI to process and synthesize varied data types are anticipated to considerably augment the analytical framework proposed herein. This paper endeavors to elucidate this hypothesis, emphasizing the need for an expanded and inclusive data analysis methodology in the realm of financial market analysis. The integration of these diverse data modalities and advanced AI technologies is not merely a response to the challenges posed by the burgeoning data volumes but also a strategic move to harness the potential embedded within this data-rich digital environment, as highlighted by the projections in Statista (2023) and the insights from Cambria et al. (2014) (Cambria and White (2014)).

2   Optimal utilization of NLP methods in finance

Automated techniques for processing financial data emerged in the 1980s Xing et al. (2018). This field has evolved greatly, alongside improvements in NLP methods. In this section, we describe methods for effectively using NLP in finance, with the goal of enhancing the precision and efficiency of financial data analysis.

2.1   Integrating diverse methodologies in NLP for advanced financial data analysis

The advent of Generative AI marks a significant milestone, particularly in its capacity to generate extensive training data for NLP. Typically, this data is multi-layered, encompassing various types and sources. Generative AI significantly simplifies these manual processes, thereby enhancing the efficiency of data interpretation. The use of NLP methods in financial forecasting has gained such traction that it is now commonly referred to as Natural Language-Based Financial Forecasting (NLFF) Xing et al. (2018). Given the interdisciplinary nature of NLP, it encompasses a blend of algorithmic techniques and principles from economics, psychology, and sociology.

2.2   Using real-time data for analysis

The popularity of social media platforms around 2010 opened new possibilities for analyzing real-time data. This feasibility was enabled by analyzing posts on these platforms, typically concise and adhering to a uniform syntax for opinion expression, a characteristic exemplified by Twitter's (currently known as X) character constraint. The stock market is often perceived as a mirror reflecting economic events. However, real-time data should not be considered the sole source of truth. The past and current prices of financial instruments may not reliably predict future prices, as suggested by Li et al. (2020). Large datasets, such as investor opinions, can be collected using web crawling technologies and stored in a format conducive to effective analysis Bi (2022). The auto-generation of financial content could be utilized for predictions about how specific events might influence the financial market. The combination of historical data with news text is also a promising approach for analysis Nikfarjam et al. (2010).

2.3   Categorizing the input from different sources

The financial market is characterized by a plethora of data sources, including financial reports (such as financial statements), corporate news, aggregated information from external bodies, and social media content. Each source provides unique information, yet they share some common features specific to their respective categories. Various methods have been developed to analyze this input, commonly focusing on aspects such as text length, objectivity, and tone. While literature offers numerous established techniques for text interpretation, the approach to non-text elements is relatively nascent. Advances in auto-interpretation technologies now allow for the inclusion of image summaries. Furthermore, the extensive array of podcasts can be exploited not only for automated transcription but also for analyzing the emotions of the speakers. Incorporating all these data types offers a more comprehensive perspective in the context of interpretability.

2.4 Advancing Financial Analysis with Multimodal Theory and Automated Multi-Modal Text Analysis

We argue that recent theorizing on multimodality Höllerer et al. (2019), combined with what we already know about the role of linguistic elements in financial markets holds the key to overcoming the known limitations of purely verbal text focused research in finance. Given the rise of various sources of communication that might affect financial market movements, we juxtapose that multimodal theory is theoretically expedient for increasing the granularity of our understanding of the movement of financial assets by highlighting the multitude of different materials and 'meaning resources' market participants invoke Jancsary et al. (2016). In broad strokes, multimodal research is predicated on the basic assumption that people in and around organizations do not produce and convey meaning solely through linguistic forms of communication, or verbal text (speech and written text), but also through visual text (images, objects, bodily movement). Verbal text is structured in a sequential and linear manner, and often uses factual statements and rational arguments to convey abstract ideas and concepts Meyer et al. (2018). This is indicative of what Clarke et al. (2019) refer to as literal language that conveys information in an unambiguous, fact-like manner. Literal expressions are often complemented in real-life settings by a type of verbal text using figurative stories or symbolic anecdotes and metaphors to relay a more imaginative sense for the respective venture. Verbal text is strongly bound in cultural conventions Höllerer et al. (2019). Visual text on the other hand conveys meaning in a more holistic and emotion-evoking manner Eisenman (2013). As Meyer et al. (2018) make clear, visual text enables a more immediate effect on comprehension than verbal text, not least because it appeals to sensory experience. Consequently, the effect of visual text tends to be quick and often has a profound influence on the receiver's emotions and attitudes Joffe (2008). Meaning conveyed by visual text tends to be more fluid and indeterminate compared to verbal text (especially literal text), which means that individuals may interpret a particular visual cue differently based on prior experience Kernbach et al. (2015). A multimodal perspective of financial asset movement marks an apt choice for this hypothesis paper, first because publicly traded companies tend to employ both verbal and visual text, and second, because information sources for those making the investment decisions such as communication videos or corporate financial communication tend to combine verbal and visual information.

3 Most effective techniques of NLP and text analysis methods

In financial market analysis, the application of Natural Language Processing (NLP) and text analysis is increasingly underscored by its mathematical foundations. This section delves into the mathematical models and algorithms that underpin the most effective NLP and text analysis techniques, emphasizing their role in computational finance.

1. **Sentiment Analysis**: This technique is typically grounded in classification algorithms. Mathematical models like Support Vector Machines (SVMs), which use hyperplanes in a high-dimensional space to classify data points, or neural network architectures like Convolutional Neural Networks (CNNs), which employ convolutional layers to process data in a grid-like topology, are common. These models are quantitatively trained on labeled datasets to categorize financial texts by sentiment.

2. **Topic Modeling**: Algorithms such as Latent Dirichlet Allocation (LDA) apply probabilistic models to discover latent topics within text corpora. LDA assumes documents are mixtures of topics, where a topic is characterized by a distribution over words. This is mathematically represented using Dirichlet distributions, providing a framework for uncovering thematic structures in financial documents.

3. **Named Entity Recognition (NER)**: NER in financial texts often utilizes sequence modeling. Algorithms like Hidden Markov Models (HMMs) or Conditional Random Fields (CRFs) apply statistical methods to label sequential data, where the state transition probabilities are learned from training data. These models are particularly effective in identifying and classifying financial entities in a sequence of text.

4. **Event Extraction**: This involves parsing and semantic analysis algorithms, where dependency parsing is often employed to construct syntactic dependency trees that represent grammatical relationships between words. Semantic Role Labeling (SRL) further identifies the predicate-argument structures, providing a mathematical framework for extracting meaningful events from financial texts.

5. **Machine Learning-Based Predictive Models**: These models often utilize regression techniques, where Ordinary Least Squares (OLS) regression or more complex algorithms like ARIMA (AutoRegressive Integrated Moving Average) for time series forecasting are applied. They mathematically model historical data to forecast future trends in financial markets.

6. **Deep Learning Techniques**: Techniques like Long Short-Term Memory networks (LSTMs), a type of Recurrent Neural Network (RNN), are used for their ability to model temporal sequences and their long-range dependencies, crucial for processing time-sensitive financial texts.

7. **Real-Time Analysis**: This involves algorithms designed for stream processing, where time-series data is analyzed in real-time using mathematical models that can update predictions and analyses as new data arrives, crucial for time-sensitive financial decisions.

The application of Natural Language Processing and text analysis in financial market analysis is increasingly characterized by sophisticated mathematical models and algorithms, ranging from statistical classification and probabilistic topic modeling to advanced deep learning architectures. The integration of Generative Artificial Intelligence holds further promise, bringing novel mathematical approaches like Generative Adversarial Networks (GANs) and Transformer models, which have the potential to revolutionize the generation and

interpretation of financial texts by learning and mimicking complex data distributions, thereby offering a more comprehensive and nuanced understanding of financial market dynamics.

Acknowledgements

List of Abbreviations

AI: Artificial Intelligence

NLP: Natural Language Processing

ML: Machine Learning

NLFF: Natural Language-Based Financial Forecasting

SVMs: Support Vector Machines

CNNs: Convolutional Neural Networks

LDA: Latent Dirichlet Allocation

NER: Named Entity Recognition

HMMs: Hidden Markov Models

CRFs: Conditional Random Fields

SRL: Semantic Role Labeling

OLS: Ordinary Least Square

ARIMA: AutoRegressive Integrated Moving Average

LSTMs: Long Short-Term Memory

RNNs: Recurrent Neural Network

GANs: Generative Adversarial Networks

References

Bi, J. (2022). Stock market prediction based on financial news text mining and investor sentiment recognition. *Mathematical Problems in Engineering 2022*.

Cambria, E. and B. White (2014). Jumping nlp curves: A review of natural language processing research [review article]. *IEEE Computational Intelligence Magazine 9*(2), 48–57.

Clarke, J. S., J. Cornelissen, and M. P. Healey (2019). Actions speak louder than words: How figurative language and gesturing in entrepreneurial pitches influences investment judgments. *Academy of Management Journal 62*, 355–360.

Eisenman, M. (2013). Understanding aesthetic innovation in the context of technological evolution. *Academy of Management Review 38*, 332–351.

Höllerer, M. A., T. J. van Leeuwen, D. Jancsary, R. E. Meyer, T. H. Andersen, and E. Vaara (2019). *Visual and Multimodal Research in Organization and Management Studies*. London: Routledge.

Jancsary, D., M. A. Höllerer, and R. E. Meyer (2016). Critical analysis of visual and multimodal texts. In R. Wodak and M. Maier (Eds.), *Methods of Critical Discourse Studies*, pp. 180–204. Sage.

Joffe, H. (2008). The power of visual material: Persuasion, emotion, and identification. *Diogenes 55*, 84–93.

Kernbach, S., M. J. Eppler, and S. Bresciani (2015). The use of visualization in the communication of business strategies: An experimental evaluation. *International Journal of Business Communication 52*(2), 164–187.

Li, X., P. Wu, and W. Wang (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of hong kong. *Information Processing & Management 57*(5), 102212.

Meyer, R. E., D. Jancsary, M. A. Höllerer, and E. Boxenbaum (2018). The role of verbal and visual text in the process of institutionalization. *Academy of Management Review 43*, 392–418.

Nikfarjam, A., E. Emadzadeh, and S. Muthaiyah (2010). Text mining approaches for stock market prediction. In *2010 The 2nd international conference on computer and automation engineering (ICCAE)*, Volume 4, pp. 256–260. IEEE.

Statista (2023). Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025 (in zettabytes).

Xing, F. Z., E. Cambria, and R. E. Welsch (2018). Natural language based financial forecasting: a survey. *Artificial Intelligence Review 50*(1), 49–73.