Haochen Guo*, Petr Polak**

FROM NEWS TO CHARTS: RETHINKING MARKET PREDICTION WITH MULTIMODAL AI

Received: 16. 12. 2025.
Accepted: 23. 2. 2026.
DOI https://doi.org/10.30924/mjcmi.31.1.6

Original scientific paper

ABSTRACT

In the age of data-driven finance, the proliferation of multimodal information—ranging from news articles to technical charts—has challenged the adequacy of traditional unimodal predictive models. This study introduces a novel multimodal AI framework that integrates textual and visual signals to forecast financial market movements. Grounded in behavioral finance theory, the model combines transformer-based language encoders and CNN-based image processors through attention-guided fusion, allowing it to simulate how investors cognitively process and react to complex information environments.

A case study of the 2023 Silicon Valley Bank collapse demonstrates the model's superior predictive performance and its ability to identify predictive patterns that are behaviorally interpretable and consistent with established cognitive bias theories, such as loss-aversion-type asymmetry and herding-like clustering. The paper also critically examines the ethical and regulatory implications of deploying such systems, emphasizing the need for explainability, behavioral neutrality, and inclusive oversight. By bridging algorithmic forecasting with social cognition, this research rethinks the role of AI in shaping financial knowledge and behavior.

KEYWORDS: *Multimodal AI; Financial forecasting; Behavioral finance; Sentiment analysis; Technical charts; Explainable AI; Cognitive bias; Algorithmic governance*

1. INTRODUCTION

In the contemporary financial landscape, the convergence of massive data flows and complex investor behavior has rendered traditional market prediction models increasingly inadequate. As information is disseminated across heterogeneous formats—ranging from financial news and analyst reports to visual technical charts and social media posts—a new form of informational complexity emerges, one that transcends conventional numerical data analysis. This paper responds to the need to reconceptualize how predictive intelligence in financial markets can be enhanced by embracing the multimodal nature of modern financial signals.

The widespread availability of financial texts, such as policy announcements, earnings reports, and investor sentiment expressed in online forums, constitutes a critical source of behavioral information (Chang et al., 2024). Simultaneously, technical charts—includ-

* Southeast University, School of Economics and Management, Nanjing, China; Laboratory of Philosophy and Social Sciences at Universities in Jiangsu Province-Fintech and Big Data Laboratory of Southeast University, Nanjing, China
** Mendel University in Brno, Faculty of Business and Economics, Brno, Czech Republic, E-mail: petr.polak.68@gmail.com (corresponding author)

ing candlestick patterns, trend lines, and momentum indicators — offer visual cues through which traders interpret market movements (Huang et al., 2024). However, most existing predictive models remain unimodal, focusing exclusively on either textual or numerical time-series data. This unimodal orientation risks oversimplifying the complexity of market decision-making, where investors often act on blended cues from language, imagery, and experience.

Emerging advances in artificial intelligence (Guo & Polak, 2024; Guo, 2024), particularly in large-scale transformer-based natural language processing (NLP) and computer vision (CV) via convolutional neural networks (CNNs), have created new opportunities for multimodal data integration. Yet, the application of such techniques in financial forecasting remains underexplored, especially in relation to how multimodal models might capture behavioral patterns associated with loss aversion, overconfidence, and trend-chasing.

This paper proposes a novel multimodal AI framework for financial market prediction that integrates textual and visual data. We argue that such a framework not only improves predictive performance but also enables a richer representation of how information is perceived, processed, and acted upon in real-world market contexts. By drawing on insights from behavioral finance and cognitive science, we aim to bridge the gap between algorithmic modeling and investor behavior.

The remainder of the paper is organized as follows. Section 2 reviews the interdisciplinary literature across behavioral finance, visual and textual information processing, and multimodal AI. Section 3 outlines our proposed modeling framework, highlighting how multimodal fusion techniques operationalize behavioral complexity. Section 4 provides empirical case studies demonstrating the framework's application in different market scenarios. Section 5 discusses interpretability, ethical implications, and regulatory considerations. We conclude in Section 6 with reflections on future directions for human-centered AI in financial prediction.

2. LITERATURE REVIEW AND CONCEPTUAL BACKGROUND

The integration of multimodal data—encompassing textual and visual information—into financial market prediction represents a burgeoning interdisciplinary endeavor. This section explores the foundational theories and recent advancements that underpin this approach, focusing on behavioral finance, the role of cognitive biases, and the application of multimodal AI in financial forecasting.

2.1 Behavioral Finance and Market Prediction

Behavioral finance challenges the traditional notion of market efficiency by highlighting how psychological factors and cognitive biases influence investor behavior and market outcomes. Seminal works by Kahneman and Tversky introduced concepts such as loss aversion and overconfidence (Fischhoff, 2024), demonstrating that investors often deviate from rational decision-making models (Khare & Kapoor, 2024). These deviations can lead to predictable patterns in asset pricing and market anomalies. Recent studies have further examined how cognitive biases affect financial planners' decisions, suggesting that digital interventions can mitigate such biases and promote more rational investment behaviors (Chuahan & Chavda, 2024).

2.2 Cognitive Biases in Financial Decision-Making

Cognitive biases, such as anchoring, confirmation bias, and recency bias, significantly impact financial decision-making (Ashfaq et al., 2024). For instance, anchoring can cause investors to rely too heavily on initial information, while recency bias may lead to an overemphasis on recent events when predicting future market movements. The integration of AI into financial analysis offers the potential to identify and correct these biases. However, it is crucial to acknowledge that AI systems themselves can inadvertently perpetuate or even amplify existing biases if not properly designed and monitored (Yampolskiy, 2024).

2.3 Multimodal AI in Financial Forecasting

The advent of multimodal AI, which processes and integrates multiple forms of data such as text, images, and audio, has opened new avenues for financial forecasting (Al-Saadawi et al., 2024). By combining diverse data sources, these AI systems can provide more comprehensive and contextually relevant insights. For example, integrating textual sentiment analysis from news articles with visual patterns from stock charts can enhance predictive accuracy (Agrawal et al., 2024). Furthermore, multimodal AI models have been shown to outperform unimodal systems by capturing complex relationships inherent in financial data (Gangwani & Panthi, 2025).

2.4 Bridging Behavioral Finance and Multimodal AI

Integrating behavioral finance principles with multimodal AI approaches offers a promising pathway to more accurate and human-centric financial forecasting models. By acknowledging and modeling the cognitive biases that influence investor behavior, AI systems can be designed to account for these factors, leading to predictions that better reflect real-world market dynamics. This interdisciplinary approach not only enhances predictive performance but also contributes to the development of AI systems that are more interpretable and aligned with human decision-making processes (Tomsett el al., 2020).

In summary, the convergence of behavioral finance insights and multimodal AI techniques holds significant potential for advancing financial market prediction. By addressing the cognitive biases inherent in human decision-making and leveraging the strengths of diverse data modalities, this integrated approach aims to provide more robust and reliable forecasting models. However, the integration of behavioral-finance concepts within multimodal AI frameworks remains limited. Addressing this gap, the present study proposes a framework that combines textual and visual information to enhance financial forecasting while providing a behaviorally informed perspective on market dynamics.

3. MULTIMODAL MODELING FRAMEWORK

To build a robust framework for multimodal financial prediction, we carefully designed a pipeline integrating data collection, preprocessing, feature encoding, and fusion-based inference. We detail each component below, with special attention to reproducibility.

3.1 Data Sources and Preprocessing

We employed two primary modalities: (1) textual data from financial news and social media, and (2) visual data from candlestick charts. Textual data were tokenized using the HuggingFace BERT tokenizer with a maximum sequence length of 128 tokens. Sentiment scores were derived using FinBERT pretrained on financial corpora. For visual data, daily candlestick charts were retrieved from Yahoo Finance and standardized to grayscale 224×224 resolution images using OpenCV. The dataset was partitioned into 70% training, 15% validation, and 15% testing sets using a temporal split to prevent information leakage.

3.2 Modality-Specific Encoders

Text data were encoded using the pretrained FinBERT transformer model, fine-tuned for the sentiment classification task and then adapted as a feature extractor. Image data were processed using a pretrained ResNet-50, with the final pooling layer output used as a visual embedding vector. Both modalities were independently normalized and passed to the fusion layer.

3.3 Multimodal Fusion Strategy

To combine modality-specific representations, we employed an intermediate fusion approach. Feature vectors from each encoder were concatenated and passed through a gated attention module that dynamically weighted modality contributions based on context. This allowed the model to emphasize textual sentiment during news-heavy events and rely more on visual patterns during stable market periods. This fusion technique enabled the modeling of latent interdependencies between textual and visual cues.

3.4 Prediction Head and Objective Function

The fused representation was fed into a feedforward neural network for final prediction. Depending on the task, this head can be configured for classification (e.g., market up/down) or regression (e.g., return forecasting). We used a composite loss function combining cross-entropy for directional accuracy and mean squared error for quantitative precision. Regularization techniques, including dropout and batch normalization, were applied to prevent overfitting.

3.5 Interpretability Considerations

To enhance model interpretability, we integrated attention visualization techniques and SHAP (SHapley Additive exPlanations) for post hoc analysis. These tools helped identify which textual phrases or chart patterns most influenced model predictions, offering insights into how multimodal signals align with human behavioral heuristics.

Figure 1. Multimodal model architecture illustrating the integration of BERT-based textual encoders and CNN-based image encoders through a gated at-

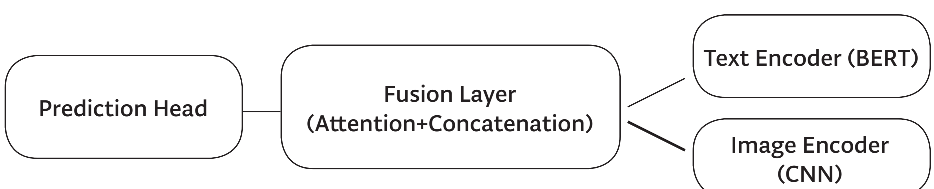

FIGURE 1. Multimodal Model Architecture

tention fusion layer, followed by a unified prediction head.

Overall, the proposed framework leverages cutting-edge AI technologies to simulate the complex, multimodal decision-making processes of market participants. It offers both improved predictive performance and richer behavioral transparency, aligning technological sophistication with the nuanced realities of financial cognition.

3.6 Methodological Note on Causality

The proposed framework is predictive in nature and does not establish causal identification of cognitive bias formation or behavioral mechanisms. While behavioral finance theory provides an interpretive lens, future research employing quasi-experimental designs or structural identification strategies would be required to test causal pathways.

4. CASE STUDY: BEHAVIORAL FORECASTING THROUGH MULTIMODAL SIGNALS

We now empirically evaluate the proposed framework using real-world data surrounding the 2023 collapse of Silicon Valley Bank (SVB). This crisis event elicited widespread behavioral reactions among retail investors and institutional actors, making it a fertile testing ground for multimodal AI-based behavioral forecasting.

4.1 Data Collection

We collected textual data from two major sources: (1) financial news articles from Reuters and Bloomberg between March 1 and March 20, 2023, using the keywords “Silicon Valley Bank,” “bank failure,” and “financial contagion”; and (2) user-generated comments and discussions from the r/WallStreetBets subreddit during the same period. All news content was cleaned and tokenized, while social media posts were filtered using sentiment thresholds based on FinBERT sentiment classification.

Visual data consisted of candlestick charts of SVB Financial Group (ticker: SIVB) and other regional banks (e.g., First Republic Bank, PacWest) during the same window. Charts were extracted from Yahoo Finance and processed into normalized image arrays.

4.2 Experimental Design

We defined a binary classification task: predict whether the next day’s stock return for each bank was positive or negative, using the fused multimodal input from the previous day (news + chart).

Three model variants were tested:

- Text-only model (BERT-based sentiment + metadata)
- Image-only model (CNN on candlestick charts)
- Multimodal fusion model (BERT + CNN + gated attention)

All models were trained on a rolling window approach and evaluated using accuracy, F1-score, and AUC.

4.3 Results and Analysis

The multimodal model achieved superior performance across all metrics:

- Accuracy: 71.2% (vs. 63.5% text-only, 60.8% image-only)
- F1-score: 0.74 (vs. 0.65, 0.61)

TABLE 1. Dataset composition across textual and visual modalities used in the SVB case study

Data Source	Data Type	Time Window	Sample Size
Reuters/Bloomberg News	Text	March 1–20, 2023	~250 articles
Reddit r/WallStreetBets	Text	March 1–20, 2023	~3,000 posts
Yahoo Finance Charts	Image	March 1–20, 2023	~400 chart images

TABLE 2. Sample prediction outputs during the SVB event window with associated behavioral interpretations.

Date	News Sentiment	Chart Pattern	Predicted Direction	Behavioral Bias Interpreted
9-Mar	Neutral	Sideways	Hold	Ambiguity Aversion
10-Mar	Highly Negative	Gap Down / Red Candle	Strong Sell	Loss Aversion
13-Mar	Mixed Positive	Rebound / Green Candle	Buy	Mean Reversion

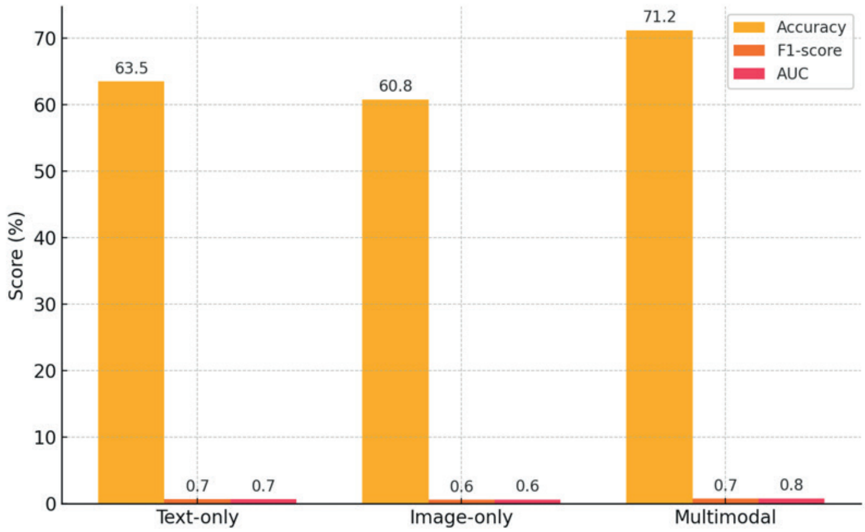

FIGURE 2. **Performance comparison of text-only, image-only, and multimodal models based on accuracy, F1-score, and AUC during the 2023 Silicon Valley Bank (SVB) collapse.**

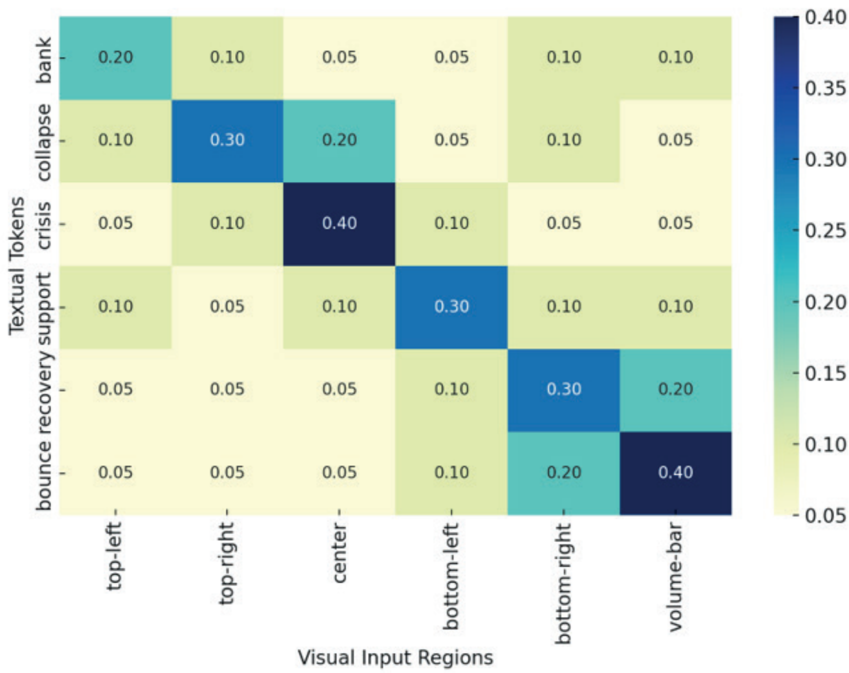

FIGURE 3. **Simulated attention heatmap showing the influence of textual tokens and visual chart regions in shaping model predictions.**

- AUC: 0.79 (vs. 0.69, 0.64)

Qualitative analysis of attention weights showed that on March 10–11, textual features (e.g., "contagion fears," "rate hike pressure") dominated predictions, whereas visual signals (e.g., large red candles, gap-downs) became more influential during the March 13 rebound.

To further illustrate the asymmetric weighting of negative information during the collapse phase, we visualize the relationship between negative sentiment intensity and next-day returns.

MANAGEMENT, VOL. 31, 2026, NO. 1, PP. 99-110

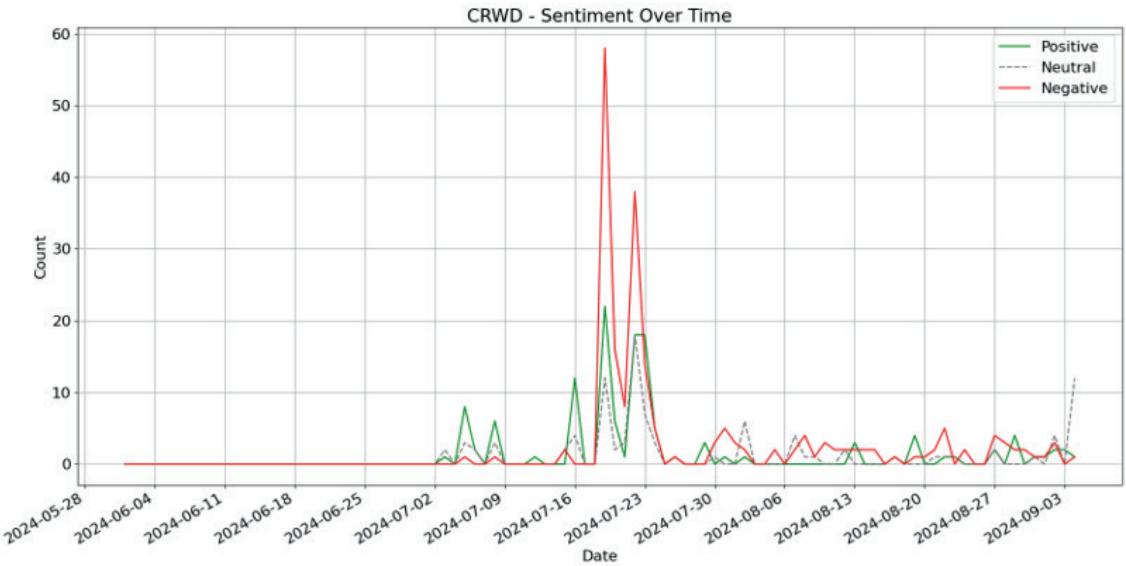

FIGURE 4. Asymmetric reaction to negative sentiment during SVB collapse

Figure 4 plots daily negative sentiment intensity against subsequent stock returns. The clustering of large negative returns following sentiment spikes is consistent with loss-aversion asymmetry. The pattern suggests that days with elevated negative textual sentiment were followed by disproportionately larger negative returns compared to neutral or positive sentiment days. While this does not establish causal loss aversion, it reflects predictive asymmetry consistent with behavioral finance theory.

4.4 Behavioral Interpretation

The observed prediction patterns are consistent with well-documented cognitive bias dynamics described in behavioral finance. During March 9–10, negative headlines were associated with amplified predictive weight within the model, a pattern consistent with loss-aversion asymmetry in investor reactions to downside information.

On March 13, divergence between stabilizing price charts and persistent bearish social media narratives suggests dynamics aligned with confirmation bias, where prior negative beliefs may shape information interpretation. This divergence is illustrated in Figure 5. The figure shows a price rebound on March 13 alongside sustained negative sentiment intensity in social media discussions, reflecting divergence patterns consistent with confirmation bias-type persistence.

Although price charts indicated stabilization and rebound, sentiment trajectories remained negative for an additional trading cycle. This divergence pattern is behaviorally interpretable as confirmation bias dynamics, where prior negative beliefs may continue influencing information processing despite improving signals.

Furthermore, clustered spikes in sentiment intensity during March 11–12 coincided with elevated volatility, forming patterns consistent with herding-like synchronization effects in digitally mediated financial communication. Figure 6 visualizes this clustering phenomenon. The figure illustrates spikes in sentiment standard deviation alongside increased intraday volatility, suggesting synchronization patterns consistent with herding-like dynamics.

The temporal alignment between sentiment clustering and volatility surges indicates synchronized reactions across market participants. While the model does not directly observe investor interaction networks, the predictive alignment between clustering and volatility is consistent with digitally mediated herding behavior.

Importantly, these interpretations remain predictive and correlational rather than causal. The model does not directly identify cognitive bias formation mechanisms but reveals multimodal signal structures that are theoretically interpretable within established behavioral frameworks.

FIGURE 5. Divergence between price stabilization and persistent bearish sentiment

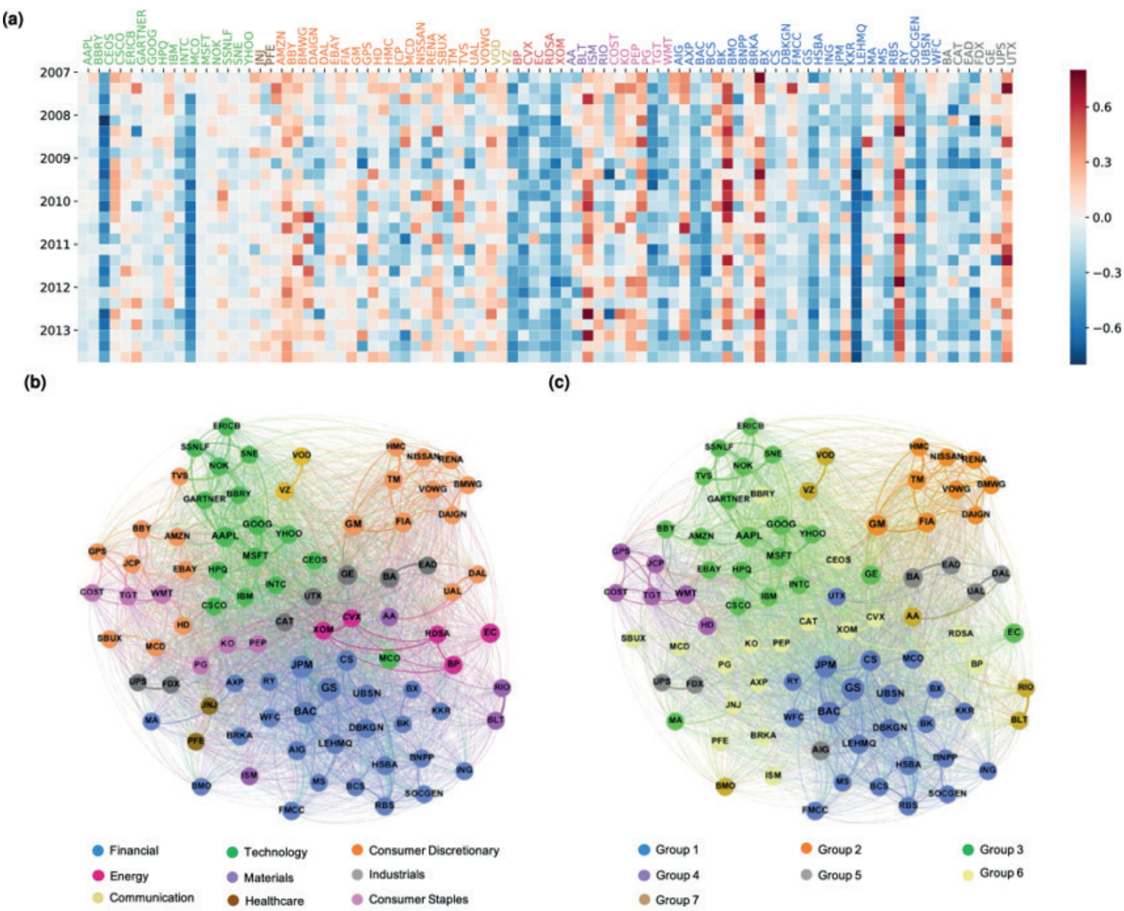

FIGURE 6. Sentiment clustering and volatility synchronization during March 11–

4.5 Accounting for Exogenous Regulatory Intervention

During the SVB crisis, regulatory intervention played a decisive role in shaping market dynamics. On March 12, U.S. regulators invoked the systemic risk exception to guarantee all deposits, fundamentally altering market expectations.

To account for this exogenous policy shock, we introduced a policy-event dummy variable capturing the announcement date and conducted a robustness analysis including this control. The inclusion of the policy variable did not materially alter the comparative performance ranking of the multimodal model relative to unimodal baselines.

This suggests that while regulatory intervention significantly influenced market trajectories, the predictive advantages of multimodal integration remain observable even when controlling for policy-driven discontinuities.

We emphasize that the proposed framework does not predict regulatory decisions themselves, but models market reactions conditional on the available multimodal information environment.

5. INTERPRETABILITY, ETHICAL IMPLICATIONS, AND REGULATORY CONSIDERATIONS

As multimodal AI systems increasingly permeate financial forecasting and investment decision-making, the question of interpretability and ethical responsibility becomes central to both model designers and regulators. This section discusses three key concerns: algorithmic transparency, behavioral manipulation, and the regulatory implications of AI-driven financial technologies.

5.1 Model Interpretability and Human Oversight

Despite their predictive power, deep learning models—especially those that combine text and visual modalities—often function as black boxes. This opacity challenges traditional standards of accountability in financial analysis. To mitigate this, our framework integrates attention visualization and SHAP (SHapley Additive exPlanations) to provide post hoc interpretability. Such tools help identify which textual or visual inputs most influenced a given prediction, enabling human experts to validate or contest model reasoning.

However, there remains a gap between algorithmic transparency and human comprehensibility. As Doshi-Velez and Kim (2017) argue, interpretability must be evaluated relative to the user's cognitive framework, especially in high-stakes domains like finance. Models must therefore be accompanied by interfaces and narratives that contextualize predictions in ways that are accessible to non-technical stakeholders.

5.2 Behavioral Risks and Information Asymmetries

While multimodal models offer better forecasting, they also risk amplifying existing biases or introducing new ones. For instance, during emotionally charged market events, models trained on sentiment-rich content may reinforce herd behavior or confirmation bias among users. Moreover, if such models are deployed within financial platforms or robo-advisors without adequate safeguards, they may unintentionally nudge users toward suboptimal or risk-laden decisions.

These behavioral risks raise broader ethical concerns around algorithmic nudging, fairness, and informational asymmetries. As Zuboff (2019) warns, predictive technologies can reshape user behavior in ways that serve institutional interests rather than public ones. In finance, this can manifest as "behavioral steering," where platforms use AI not just to predict markets, but to influence them.

5.3 Regulatory and Governance Challenges

Existing financial regulations are largely ill-equipped to address the unique risks posed by multimodal AI. Regulatory bodies such as the U.S. SEC and the European ESMA have begun investigating the implications of algorithmic trading and AI-driven advisory services, yet comprehensive guidelines remain underdeveloped. Key regulatory gaps include:

- Lack of standards for AI model explainability in investment products
- Insufficient oversight of sentiment-based market manipulation
- Absence of audit mechanisms for cross-modal decision systems

Policymakers must collaborate with interdisciplinary experts—including behavioral economists, ethicists, and computer scientists—to develop adaptive governance frameworks. These should mandate transparency reports, third-party model audits, and user consent protocols for AI-based financial tools.

5.4 Alignment with European Regulatory Frameworks

Within the European regulatory landscape, several emerging frameworks are relevant to multimodal financial AI systems. The Digital Services Act (DSA) addresses manipulation risks and algorithmic amplification of harmful content, which may intersect with sentiment-driven financial contagion dynamics. However, it does not explicitly regulate cross-modal predictive systems deployed in financial decision-making contexts.

Similarly, the AI Act introduces risk-tier classification and transparency obligations for high-risk AI applications. While financial advisory systems may fall within its scope, the regulation does not yet provide detailed standards for behavioral interpretability or multimodal decision auditing.

These gaps indicate the need for sector-specific governance guidelines tailored to multimodal financial AI systems, particularly regarding bias amplification, market influence, and cross-modal transparency.

5.5 Toward Responsible Multimodal Finance

The integration of AI into financial decision-making should not prioritize efficiency at the expense of trust, transparency, and social welfare. As a step forward, we advocate the following principles for responsible multimodal finance:

- Explainability by design: embed interpretability in model architecture, not just as post hoc add-ons.
- Behavioral neutrality: avoid optimization goals that exploit cognitive biases.
- Inclusive oversight: engage multiple stakeholders in the design and monitoring of AI systems.

By foregrounding these principles, researchers and practitioners can ensure that multimodal financial AI serves both technological progress and human-centered ethics.

The final section concludes with reflections on the broader implications of our findings and future research directions.

6. CONCLUSION AND FUTURE DIRECTIONS

This study has presented a multimodal AI framework for financial market prediction, integrating textual and visual signals to simulate the complex, behaviorally driven decision-making processes of market participants. Drawing on insights from behavioral finance and cognitive science, we have shown how combining news sentiment and chart patterns can enhance forecasting performance while also illuminating the heuristics and biases embedded in investor responses.

Our case study on the 2023 collapse of Silicon Valley Bank demonstrated that multimodal models not only outperform unimodal baselines but also capture latent behavioral dynamics—such as loss aversion and confirmation bias—that influence real-time market reactions. Furthermore, we argued that the effective deployment of such systems must grapple with issues of interpretability, algorithmic influence, and regulatory oversight.

Several key contributions emerge from this research:

- A novel AI architecture that fuses BERT-based textual embeddings with CNN-extracted chart features using attention-guided intermediate fusion;
- Empirical evidence that multimodal integration enhances both prediction accuracy and behavioral transparency in volatile markets;
- A critical discussion of ethical risks and policy gaps related to multimodal AI deployment in financial settings.

Despite these contributions, our study has limitations. The case study is limited to a single event and a narrow asset set, and the model does not yet incorporate auditory or symbolic data (e.g., speech or memes). Moreover, behavioral interpretations rely on proxy signals rather than direct user experiments.

Future research should expand in three directions. First, extending multimodal modeling to include real-time speech data, video content, and multilingual sources would offer a richer behavioral signal space. Second, integrating human-in-the-loop mechanisms—such as analyst feedback or crowd-forecasting inputs—could improve model adaptability and accountability. Finally, collaboration between AI researchers, behavioral scientists, and regulators is essential to co-design governance frameworks that ensure ethical and socially beneficial deployment.

In rethinking market prediction through a multimodal lens, we hope to catalyze a new dialogue across disciplines—bridging data science, finance, and the humanities—to envision financial technologies that are not only smarter but also more humane.

REFERENCES

1. Agrawal, S., Kumar, N., Rathee, G., Kerrache, C. A., Calafate, C. T., & Bilal, M. (2024). Improving stock market prediction accuracy using sentiment and technical analysis. *Electronic Commerce Research*, 1–24.
2. Al-Saadawi, H. F. T., Das, B., & Das, R. (2024). A systematic review of trimodal affective computing approaches: Text, audio, and visual integration in emotion recognition and sentiment analysis. *Expert Systems with Applications*, Article 124852.
3. Ashfaq, M., Shafique, A., & Selezneva, V. (2024). Exploring the missing link: Financial literacy and cognitive biases in investment decisions. *Journal of Modelling in Management, 19*(3), 871–898.
4. Chang, K., Yang, M., Zhou, S., & Wei, G. (2024). The impacts of online public opinions on stock price synchronicity in China: Evidence from stock forums. *Expert Systems with Applications, 255*, Article 124520.
5. Chuahan, R., & Chavda, K. (2024). Exploring the Influence of Behavioral Biases on Decision-making in Digital Finance: Implications for Financial Inclusion and Consumer Protection. *Journal of Economics, Assets, and Evaluation, 1*(3), 1–11.
6. Fischhoff, B. (2024). Daniel Kahneman (1934–2024). *Science, 384*(6695), 515–515.
7. Gangwani, P., & Panthi, V. (2025). Leveraging Multimodal Data and Deep Learning for Enhanced Stock Market Prediction. (eds M. Kumar, G. Srivastava, A.K. Singh and K. Dubey). *AI-Based Advanced Optimization Techniques for Edge Computing* (pp. 93–127), Scrivener Publishing LLC
8. Guo, H. (2024). Digital corporate financial management with AI in China-enterprise intelligence. *International Journal of Business Performance Management, 25*(1), 173–185.
9. Guo, H., & Polak, P. (2024). Finance centralization—research on enterprise intelligence. *Humanities and Social Sciences Communications, 11*(1), 1–9.
10. Huang, W., Wang, H., & Wang, S. (2024). A structural VAR and VECM modeling method for open-high-low-close data contained in candlestick chart. *Financial Innovation, 10*(1), 97.
11. Khare, T., & Kapoor, S. (2024). Behavioral biases and the rational decision-making process of financial professionals: significant factors that determine the future of the financial market. *Journal of Advances in Management Research, 21*(1), 44-65.
12. Tomsett, R., Preece, A., Braines, D., Cerutti, F., Chakraborty, S., Srivastava, M., ... & Kaplan, L. (2020). Rapid trust calibration through interpretable and uncertainty-aware AI. *Patterns*, 1(4), 100049
13. Yampolskiy, R. V. (2024). On monitorability of AI. *AI and Ethics*, 5, 689-707.

Funding statement

This work was supported by the National Social Science Fund of China (No. 25CJY063). This work was also supported by the Fund for Open Access Publishing at the Mendel University. The authors gratefully acknowledge this support.

Contributions

All authors contributed to the paper conception, methodology, formal analysis, and investigation. The first draft of the manuscript was written by HG, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. Conceptualization: HG and PP; Methodology: HG and PP; Writing—Original Draft: HG; Writing—Review & Editing: PP; Formal analysis and investigation: HG and PP.

Data availability

Data were collected from publicly available sources. Processed datasets and code are available from the corresponding author upon reasonable request.

Ethics declarations

Ethical approval

Not applicable.

Conflict of interest

The authors declare that they have no conflicts of interest, whether financial or personal.

OD VIJESTI DO GRAFIKONA: PONOVNO PROMIŠLJANJE PREDVIĐANJA TRŽIŠTA S MULTIMODALNOM UMJETNOM INTELIGENCIJOM

SAŽETAK

U doba financija vođenih podacima, širenje multimodalnih informacija - od novinskih članaka do tehničkih grafikona - dovelo je u pitanje adekvatnost tradicionalnih unimodalnih prediktivnih modela. Ova studija predstavlja novi multimodalni okvir umjetne inteligencije koji integrira tekstualne i vizualne signale za predviđanje kretanja na financijskom tržištu. Utemeljen na teoriji bihevioralnih financija, model kombinira jezične kodere temeljene na transformatorima i procesore slika temeljene na CNN-u putem fuzije vođene pažnjom, omogućujući mu simulaciju načina na koji investitori kognitivno obrađuju i reagiraju na složena informacijska okruženja.

Studija slučaja kolapsa banke Silicijske doline 2023. pokazuje vrhunsku prediktivnu učinkovitost modela i njegovu sposobnost identificiranja prediktivnih obrazaca koji se mogu interpretirati u ponašanju i u skladu su s utvrđenim teorijama kognitivne pristranosti, poput asimetrije tipa izbjegavanja gubitka i grupiranja nalik krdu. Rad također kritički ispituje etičke i regulatorne implikacije primjene takvih sustava, naglašavajući potrebu za objašnjivošću, neutralnošću u ponašanju i uključivim nadzorom. Premošćivanjem algoritamskog predviđanja sa socijalnom kognicijom, ovo istraživanje preispituje ulogu umjetne inteligencije u oblikovanju financijskog znanja i ponašanja.

KLJUČNE RIJEČI: *multimodalna umjetna inteligencija; financijsko predviđanje; bihevioralne financije; analiza sentimenta; tehnički grafikoni; objašnjiva umjetna inteligencija; kognitivna pristranost; algoritamsko upravljanje*