Tracing Knowledge Flows in Financial Markets with Hybrid Sentiment AI

Peter Németh and Mária Bohdalová
Comenius University Bratislava, Bratislava, Slovakia

nemeth83@uniba.sk
maria.bohdalova@fm.uniba.sk

Abstract: Financial markets are commonly described as information-efficient, yet the phrase conceals more complexity than it clarifies. Between the arrival of information and the adjustment of prices lies a process that is neither fully observable nor mechanically uniform. Interpretation mediates how signals are received, credibility shapes their perceived relevance, and responses unfold unevenly across participants. From a knowledge management perspective, this raises a central question: how does dispersed information consolidate into knowledge that meaningfully informs decisions? The framework we developed does not seek to model the entire market mechanism. We concentrate instead on a measurable segment of that transformation. Media sentiment is examined using a hybrid analytical approach that combines the lexicon-based VADER model with the domain-specific transformer model FinBERT. The two models do not always converge, and this divergence provides additional insight rather than methodological weakness. Their integration enables us to capture both surface polarity and contextual financial nuance within a single structure. Because informational inputs differ in reliability, sources are weighted by credibility and filtered through an entity-based relevance mechanism. Sentiment scores are aggregated across adjustable time windows, enabling us to detect whether informational pressure accumulates before a price adjustment. Using selected S&P 500 firms characterised by sustained media exposure and high liquidity as the empirical setting, we conduct structured backtesting to evaluate directional alignment and interval coherence across multiple horizons. The main objective of this paper is to assess whether consistent relational patterns can be identified. Our results indicate that such patterns are present, although their strength varies across horizons and informational density. In some instances, aggregated sentiment precedes adjustment, while in others it dissipates without measurable impact. This unevenness suggests that sentiment analysis may be more appropriately understood as a mechanism for tracing knowledge flows within financial ecosystems rather than as a deterministic forecasting device. From this perspective, our study contributes to knowledge management by offering a computational approach for examining how dispersed information becomes actionable knowledge in high-velocity environments.

Keywords: FinBERT, Strategic forecasting, Hybrid AI framework, Financial sentiment Analysis, Distributed knowledge systems

1. Introduction

Financial markets are traditionally viewed through the lens of the Efficient Market Hypothesis (EMH), which holds that prices reflect all available information immediately. However, from the perspective of Knowledge Management (KM), this view oversimplifies a much more complex cognitive and technical process. The path from the generation of raw, scattered information to the actual adjustment of asset prices is mediated by interpretation, the credibility of sources, and varying speeds of information dissemination. In a highly dynamic environment, the main challenge is no longer a lack of data but rather information overload and significant semantic noise.

The central question this study addresses is the mechanisms by which unstructured news data is consolidated into actionable market insights. Although sentiment analysis has become a standard tool for financial forecasting, traditional models often lack the sensitivity needed to process complex financial narratives or the varying reliability of news sources. A gap between signal and noise arises, where meaningful insights are often diluted by general market noise or fleeting media optimism. Therefore, the primary objective of this study is to assess whether aggregated media sentiment contains a measurable information signal regarding subsequent price movements in financial markets and, under what conditions, such signals are most reliable. By combining domain-specific NLP models with credibility-weighted aggregation and structured backtesting, this research advances the understanding of how dispersed financial information consolidates into actionable market knowledge.

To address these challenges, this work presents a hybrid sentiment modelling framework. Unlike homogeneous approaches, our framework employs a dual-engine logic that combines the emotional sensitivity of the VADER lexicon with the deep contextual understanding of the FinBERT transformer. At the same time, we acknowledge that the market processes information with varying time delays. Therefore, our system analyses news at configurable intervals. For purposes of this study, we tested three specific historical windows of 2, 5, and 7 days to examine the depth of historical data needed to generate a reliable sentiment signal. The model tracks how accumulated news gradually transforms into actionable knowledge within the S&P 500 ecosystem by weighing sources according to their credibility (W_s).

Overall, this study contributes to the field of knowledge management by demonstrating that computational sentiment analysis can serve as a viable mechanism for tracking how dispersed financial information consolidates into actionable market knowledge, particularly within short forecasting horizons and under conditions of sustained media exposure.

The literature review is presented in the next section. Section 3 describes our proposed methodology and the analysed data. Our results are presented in section four. Finally, section 5 concludes our findings.

2. Literature Overview

2.1 Knowledge Management and Financial Intelligence

The traditional understanding of financial markets is increasingly shifting from the analysis of numbers to the knowledge management (KM) flows. In his work, Ajayi (2022) emphasises that KM is not merely an archive of information but a critical strategy for the survival of stockbrokers, enabling the transformation of scattered investment insights into a competitive advantage. However, this process faces challenges with big data. Chierici et al. (2018) argue that the mere existence of large volumes of data does not create value. What is crucial is the specific KM practice that can transform this unstructured data into actionable knowledge. In the current context of 2026, this process cannot proceed without advanced artificial intelligence (AI). Abboud (2025) notes that modern AI systems function as strategic assets, making financial expertise accessible to the broader public by continuously scanning vast data streams and identifying warning signs that would otherwise remain hidden from human analysts. This perspective situates sentiment analysis not as a predictive algorithm but as a knowledge-extraction mechanism. This framing is consistent with KM literature's emphasis on converting raw signals into structured, actionable intelligence (Chierici et al., 2018).

2.2 Market Efficiency and the Role of Information

The theoretical basis for examining the relationship between sentiment and prices stems from Fama's (1970) Efficient Market hypothesis (EMH), which assumes that asset prices fully and immediately reflect all available information. Although the weak and semi-strong forms of the EMH have received considerable empirical support, this hypothesis is based on an idealised assumption. In practice, interpretation influences how signals are received, and differences in cognitive abilities, access speed, and assessments of credibility mean that market participants do not react uniformly. Fama (1991) himself acknowledged that evidence of the predictability of returns based on variables complicates the original framework, suggesting that not all information is immediately and fully reflected in prices. This theoretical gap, between the arrival of information and the adjustment of prices, is precisely the space in which sentiment analysis operates. It is also the space this study seeks to empirically investigate.

2.3 From Lexicons to Transformers: The Evolution of Financial Sentiment Analysis

The use of sentiment analysis as a financial intelligence tool has evolved significantly over the past two decades. Earlier approaches relied on lexicon-based models that assigned polarity scores to words in financial texts. Although these methods were readily available, they were found to have limited ability to handle the contextual ambiguity characteristic of financial language. Mishev et al. (2020) conducted a systematic comparative study evaluating a wide range of NLP approaches for financial sentiment analysis, ranging from dictionary-based methods to word- and sentence-level encoders and modern transformer architectures. Their results confirm that contextualised transformer inputs significantly outperform approaches based on dictionaries and fixed-word encoders, with the BART model achieving the highest classification accuracy among all evaluated models. Araci (2019) directly addressed the limitations of general sentiment models by introducing FinBERT, a domain-specific transformer pre-trained on financial news corpora, which substantially outperformed both lexical methods and general deep-learning classifiers on financial sentiment benchmark datasets. Buckman et al. (2020) further found, based on a sentiment index of daily news reports over a period of four decades, that sentiment indicators derived from news reports lead survey-based measures of consumer confidence by several weeks, a finding with direct implications for the timeliness of any sentiment-based signal. Based on this, Fazlija and Harder (2022) demonstrated that incorporating sentiment from FinBERT, along with historical price data, into predictive models improves the accuracy of forecasts of the S&P 500 index's direction. Puh and Bagić Babac (2023) confirmed that combining sentiment analysis with the FinBERT model and recurrent neural network architectures yields more accurate stock price forecasts than models that rely solely on historical prices, thereby strengthening the case for hybrid approaches.

2.4 Source Credibility, Aggregation Windows, and Knowledge Flow Dynamics

In addition to the choice of model, the structure of information aggregation also plays a decisive role in determining whether media signals translate into meaningful market insights. Financial news is not homogeneous, as institutional sources such as news agencies operate under different editorial standards and reach different audiences than aggregators for retail investors or social media platforms. Chierici et al. (2018) argue in the knowledge management literature that not all data gathered from social media have the same level of informativeness. Firms must therefore evaluate which inputs are worth processing through knowledge management practices to generate actionable knowledge. The time dimension is equally important, as sentiment does not immediately translate into price adjustments. Fazlija and Harder (2022) demonstrated that incorporating FinBERT-based sentiment scores from financial news improves the accuracy of S&P 500 price direction forecasts. Their results further showed that sentiment derived from content significantly outperforms that derived from headlines, suggesting that the depth of processed information plays a decisive role in the quality of the signal. Taken together, these findings support a framework in which credibility-weighted sentiment aggregation and time windows serve as a mechanism for tracking knowledge flows in financial ecosystems. This perspective is consistent with Abboud (2025), who argues that AI-driven financial systems are most valuable not as deterministic forecasting tools, but as instruments for mapping how informational pressure accumulates and dissipates over time.

While Fazlija and Harder (2022) demonstrated the predictive value of FinBERT-based sentiment, their framework does not incorporate source credibility weighting or examine how aggregation depth affects signal stability. Similarly, existing knowledge management frameworks, such as those of Chierici et al. (2018), address the transformation of big data into actionable knowledge but do not operationalise this process in a financial sentiment context. The present study addresses this gap by integrating credibility-weighted aggregation, hybrid sentiment modelling, and structured backtesting into a unified framework explicitly positioned within the knowledge management tradition.

3. Methodology

The methodology focuses on the technical design of data processing, sentiment analysis, and price trend integration. The system analyses news at configurable time intervals. Three aggregation windows were selected to capture distinct informational dynamics. The 2-day window reflects immediate, recency-weighted sentiment, capturing the most current informational signals before a price adjustment. The 5-day window corresponds to a standard trading week and reflects the accumulation of sentiment across a full cycle of market activity. The 7-day window extends coverage to a full calendar week, including weekend news and commentary that may influence market-opening conditions. Together, these windows enable a systematic comparison of how the depth of sentiment accumulation affects predictive reliability. The research implements a modular computational pipeline designed to transform unstructured financial news into quantified signals, prioritising transparency in the logic of data cleaning and sentiment weighting.

3.1 Data Acquisition and Credibility Weighting

The system utilises an automated data extraction module based on the feedparser library (McKee, 2023) to aggregate news from global financial RSS feeds. Unlike general data scrapers, this system applies a Source weight (W_s) based on perceived institutional credibility at the point of entry, following the principle that not all data inputs carry equal informative value (Chierici et al., 2018). This credibility-weighted filtering represents the first stage of knowledge refinement, in which raw informational inputs are evaluated and prioritised before entering the sentiment processing pipeline. Each news item *I* from source *i* is stored with its associated weight, ensuring that high-credibility institutional nodes have a greater influence on the resulting aggregation than broader retail aggregators.

The mathematical representation of the initial information base at time *t* is:

$$E_t = \sum_{i=1}^{n} \left(I_{i,t} \cdot W_{c,i} \right) \tag{1}$$

The weights are distributed across three tiers based on professional reputation and market relevance:

- **Tier 1 ($W_c = 1.0 - 0.9$):** Primary news agencies (e.g., Reuters, Bloomberg).
- **Tier 2 ($W_c = 0.8 - 0.7$):** Major financial news agencies (e.g., CNBC, BBC).
- **Tier 3 ($W_c = 0.6 - 0.4$):** General aggregators (e.g., Yahoo Finance, Google News).

Before the sentiment analysis, the source HTML content is processed with the newspaper3k library (Ou-Yang, 2013) for in-depth text cleaning. To maintain a high signal-to-noise ratio, the system performs three sequential layers of heuristic filtering:

1. **Structural Validation (Ticker Matching):** The algorithm scans the URL and the headline for the presence of a specific "Cashtag" (e.g., $AAPL) or an official company name. This ensures the article concerns a specific subject, not just a general overview of the sector.
2. **Frequency Analysis and Entity Significance:** The system performs a full-text analysis to calculate a composite relevance score (*R*) based on multiple signals, including the presence of the ticker or company name in the headline and URL, the frequency of entity mentions throughout the full text, and penalties for articles dominated by general market terminology without a specific company identifier. Articles where the entity appears only marginally are excluded entirely, preventing signal dilution and ensuring that sentiment reflects only knowledge flows directly relevant to the target firm.
3. **Automatic Content Filtering using regular expressions:** The system automatically extracts junk content by utilising regular expressions. This includes automated industry reports, general CAGR forecasts, and promotional summaries of research reports that lack new information based on current events.

3.2 Hybrid Sentiment Modelling (VADER + FinBERT)

A key component of the framework is a length-dependent hybrid sentiment engine. It recognises that different NLP architectures excel at different text densities and semantic complexities, a distinction confirmed by comparative evaluations of sentiment methods in financial contexts (Mishev et al., 2020). This length-dependent weighting reflects a sensemaking logic, in which the depth of textual analysis is adapted to the informational density of the input, enabling the framework to extract more nuanced knowledge signals from complex financial narratives.

A. Model convergence logic

The engine combines VADER (Valence Aware Dictionary and sEntiment Reasoner) (Hutto and Gilbert, 2014) with FinBERT, a transformer model pre-trained on financial text (Araci, 2019), which itself builds on the foundational BERT architecture (Devlin et al., 2019). Integration is governed by a length factor (*L*):

- **Headline Polarity (*L* < 40 words):** VADER is assigned a 70% weight, as short headlines rely primarily on surface-level polarity signals where lexicon-based scoring demonstrates greater sensitivity than contextual transformer models.
- **Contextual Analysis (*L* > 180 words):** FinBERT is assigned a weight of 70%, thereby leveraging its ability to decode complex financial semantics and specialised terminology.
- **Balanced weighting ($40 \leq L \leq 180$ words):** Both models contribute equally to the final score.

B. Contrastive sentence splitting

To handle the concession-signal structure common in financial reporting (e.g., profits stagnated, but growth is accelerating), the system implements a contrast multiplier. The algorithm splits sentences at contrastive conjunctions to ensure accurate capture of each clause's sentiment. The clause following the conjunction is assigned a multiplier of 1.6x (β):

$$S_{sent} = \frac{s_{pre} + (s_{post} \cdot \beta)}{1 + \beta} \tag{2}$$

3.3 Aggregation and Signal Construction

Once individual article scores are calculated, they are aggregated within adjustable time windows. The system calculates a weighted average to prevent volume-based distortion, consistent with the principle that the quality of aggregated knowledge inputs matters more than their quantity (Chierici et al., 2018):

$$S_{avg} = \frac{\sum\left(S_{core_j} \cdot W_{c,j}\right)}{\sum W_{c,j}} \quad (3)$$

This aggregation process operationalises the concept of knowledge flows, transforming dispersed media signals into a consolidated informational input that reflects the cumulative sentiment pressure within a defined time window. The resulting S_{avg} is then mapped to directional recommendations:

- **Buy ($S \geq 0.2$):** Indicates a strong positive knowledge flow, suggesting upward price pressure.
- **Neutral ($-0.2 < S < 0.2$):** Reflects a state of informational equilibrium or high signal variance where no clear directional trend is identifiable.
- **Sell ($S \leq -0.2$):** Indicates a dominant negative sentiment, indicating potential downward movement.

This directional signal represents the conversion of aggregated sentiment into actionable knowledge, providing a structured basis for investment decision-making (Chierici et al., 2018). The threshold of $\pm$0,2 was deliberately conservative relative to standard VADER classification thresholds (±0,05; Hutto and Gilbert, 2014) to reduce signal volatility and restrict directional recommendations to periods of stronger, more sustained sentiment pressure. These thresholds provide a consistent quantitative basis for the subsequent backtesting phase, enabling the predictive accuracy of the model to be thoroughly evaluated based on historical price movements.

3.4 Trend Adjustment and Forecasting Horizons

The final stage integrates real-time price data retrieved by the yfinance library (Aroussi, 2024) to calibrate the sentiment signal against market reality through two technical mechanisms:

- **The Trend Safety Mechanism:** Using the yfinance API, the system monitors price movements (T) over a 3-day lookback period. If a price decline below ($T < -1\%$) is detected, the system reduces the positive sentiment score. If the decline exceeds ($T < -2\%$), the safety mechanism is fully activated and automatically caps positive media sentiment to a neutral level. This reduces the impact of temporary optimism in news reports during periods of demonstrable asset price declines, addressing the known tendency of media sentiment to lag behind the actual market movements (Buckman et al., 2020).
- **Interpolated Predictive Horizons:** The model generates a 14-day Prediction Band that provides a detailed forecast for each upcoming trading day. Based on sentiment intensity, the model calculates daily price milestones (P_n) for the entire duration ($n = 1,2, \dots, 14$) using a sentiment-scaled interpolation scheme. The base price movement is derived from the aggregated sentiment score, which is then scaled using predefined multiplier points and linearly interpolated for intermediate days. The projected price for a given day n is determined according to the following formula:

$$P_n = P_{current} \cdot \left(1 + direction \cdot \left(M_{base} \cdot interp(n)\right)\right) \quad (4)$$

This process generates a dynamic range between values: P_{min} and P_{max} for each day in the horizon. These values are also used during the backtesting phase to assess the model's interval coherence, measuring the extent to which the actual market price consistently remains within the predicted trajectory (Fazlija and Harder, 2022).

3.5 Sample and Empirical Design

The empirical analysis was conducted on a sample of firms selected from the S&P 500 index, chosen based on sustained media exposure and high market liquidity. Data collection was performed using an automated RSS aggregation pipeline drawing from seven sources: Reuters, Bloomberg, CNBC, BBC, Yahoo Finance, Investopedia, and Google News. The observation period spans from November 2025 to March 2026, covering approximately five months of systematic data collection. A multi-layer data cleaning pipeline was applied, including blacklist filtering of generic industry reports, ticker-based relevance scoring, entity frequency analysis, and duplicate removal. A total of 10,725 directional predictions were generated and evaluated across three aggregation windows of 2, 5, and 7 days. Backtesting was conducted under the assumption of no transaction costs, with predictions evaluated against actual closing prices retrieved via the yfinance library (Aroussi, 2024).

4. Results and Discussion

4.1 Directional Accuracy Analysis

The backtesting evaluation was performed on 10,725 predictions. These were divided into three news collection windows: 2-day (n = 3,118), 5-day (n = 3,242), and 7-day (n = 4,365).

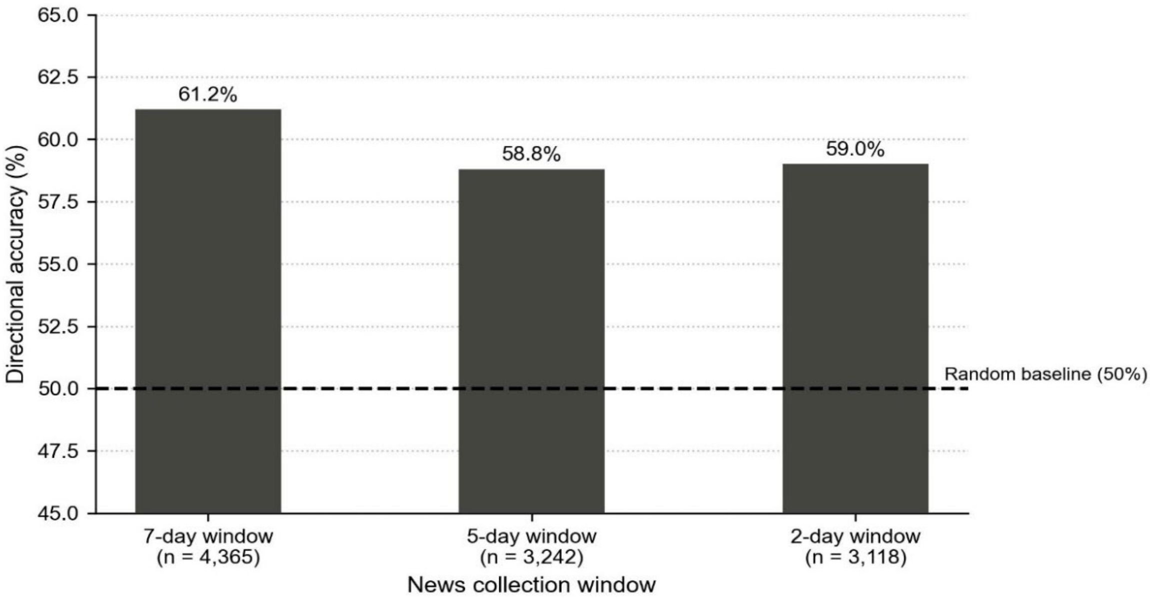

Figure 1: Overall Directional Accuracy (%) by News Collection Window. The chart compares prediction accuracy across three aggregation windows (2-day, n = 3,118; 5-day, n = 3,242; 7-day, n = 4,365) against a random baseline of 50%. All three windows exceed the baseline, with the 7-day window achieving the highest accuracy of 61,2%. Total predictions: n = 10,725. Source: Authors

As shown in Figure 1, all three windows achieved directional accuracy above the random baseline 50%. The 7-day window performed best with an accuracy of 61,2%. The 2-day and 5-day windows performed at 59,0% and 58,8%, respectively. This confirms that aggregated media sentiment carries a measurable informational signal regarding subsequent price movements. The dominance of the 7-day window suggests that broader news accumulation yields a more stable sentiment-based signal. This is consistent with Buckman et al. (2020), who found that information pressure requires time to consolidate before it generates a measurable market effect. Markets react only when sufficient informational pressure from credible sources accumulates, which a 7-day period captures more effectively than shorter windows (Ajayi, 2022; Chierici et al., 2018).

4.2 Signal Type Analysis

Figure 2 shows the accuracy distributed by signal type. The results indicate a significant asymmetry across the three recommendation categories.

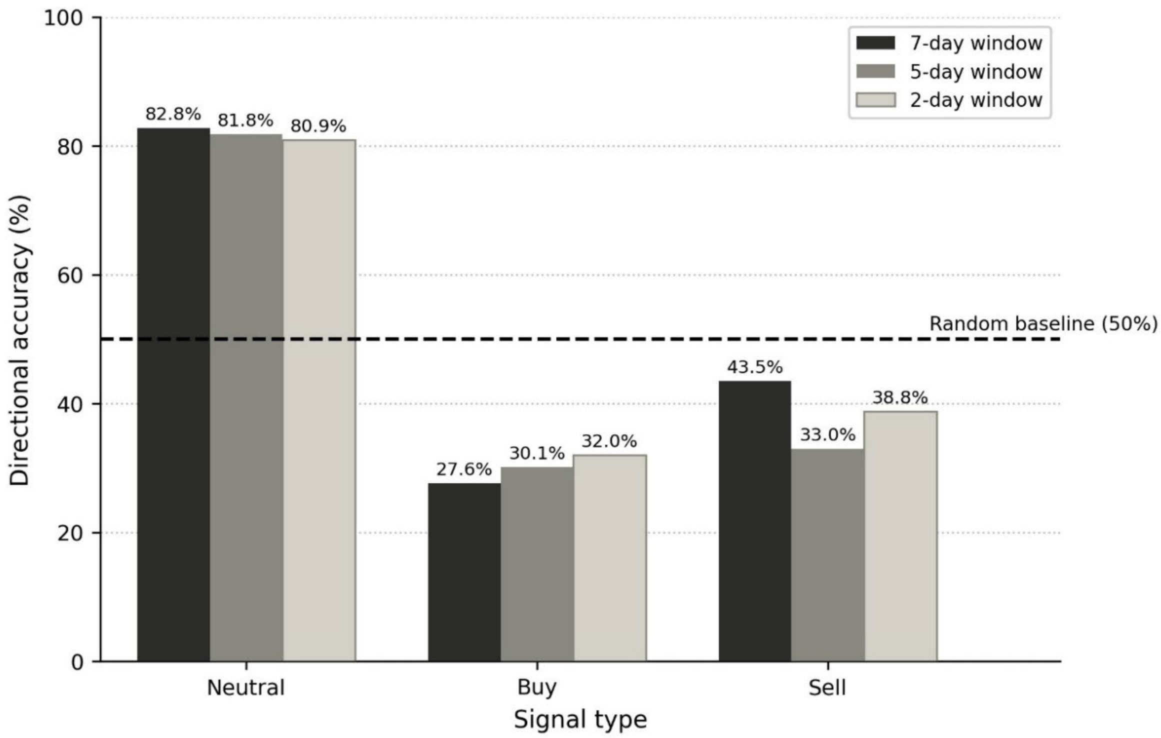

Figure 2: Directional accuracy (%) by signal type across news collection windows. The chart displays the accuracy of Buy, Neutral, and Sell signals separately for each aggregation window. The Neutral signal consistently exceeds 80% across all windows, while Buy and Sell signals fall below the random baseline of 50%. Source: Authors

The Neutral signal, which means hold, was the most accurate across windows. It reached 82,8% in the 7-day window, 81,8% in the 5-day window, and 80,9% in the 2-day window. The Buy signal fell below the random benchmark in all three windows. Accuracy ranged from 27,6% to 32,0%. The Sell signal showed average performance ranging from 33,0% to 43,5%. The consistently high accuracy of the Neutral signal warrants closer examination. From a theoretical perspective, this finding is consistent with the semi-strong form of the Efficient Market Hypothesis (Fama, 1970), which implies that asset prices already incorporate publicly available information during periods of informational equilibrium. Under such conditions, accumulated media sentiment does not generate sufficient directional pressure to produce a measurable price adjustment, a state that the framework captures more consistently than directional market reversals. It should be noted, however, that the dominance of the Neutral signal is also partially attributable to the conservative threshold of $\pm 0{,}2$ applied to the sentiment scale of -1 to +1. This boundary defines a relatively wide neutral zone, which structurally limits the frequency of Buy and Sell signals. The weak performance of directional signals should therefore be interpreted with this design choice in mind. Whether narrower or adaptive thresholds would improve the framework's discriminatory power without sacrificing stability remains an open question and represents a natural direction for future research.

4.3 Forecasting Horizon Analysis

Figure 3 illustrates how accuracy varies across different forecast horizons from day 1 to day 7. Several patterns can be observed.

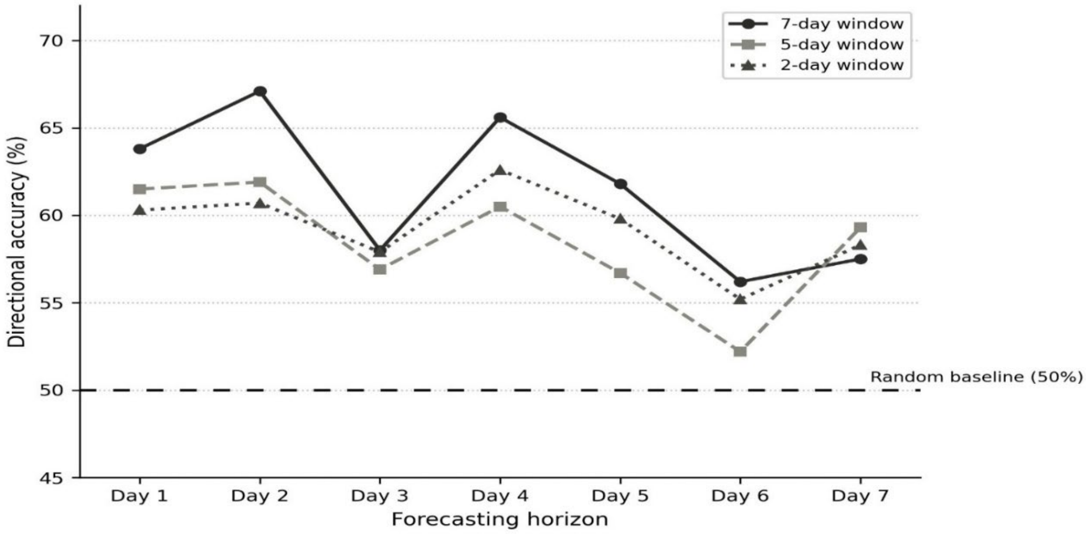

Figure 3: Directional accuracy (%) by forecasting horizon (Day 1-Day 7) across three news collection windows (2-day, 5-day, 7-day). The highest accuracy of 67,1% is achieved at the 2-day forecasting horizon with a 7-day window. All windows show a consistent decline in accuracy at the 3-day horizon. The random baseline of 50% is shown for reference. Source: Authors

First, the 7-day window has the highest accuracy at the 2-day horizon, at 67,1%. This is the highest value in the study. Second, all three windows show a consistent decline in accuracy at the 3-day horizon. This suggests that sentiment signals in this interval are experiencing a temporary dispersion effect. Third, accuracy partially recovers on the 4th day before declining more significantly on the 6th day. The 5-day window drops to 52,2% on the 6th day, approaching the random baseline. Our findings confirm that sentiment signals are most reliable within short forecasting horizons of one to two days. This is consistent with Buckman et al. (2020), who demonstrated that sentiment indicators derived from news articles capture changes in economic conditions earlier than traditional survey-based measurements, suggesting that their informational value is concentrated in the immediate period following publication. The nonlinear pattern further supports the view that sentiment functions more as an indicator of information flow than as a stable predictive variable. This pattern may reflect the multi-stage nature of information dissemination, where market participants process news at different speeds.

4.4 Interval Coherence Analysis

The interval coherence analysis evaluated what proportion of actual prices fell within the price ranges predicted by the system. Table 1 presents the results of all three time windows and forecast horizons.

Table 1: Interval coherence rates (%) by forecasting horizon (Day 1-Day 8) and news collection window. Interval coherence measures the proportion of actual closing prices that fell within the model's predicted price band for a given day. Values above 50% indicate that the prediction band captures actual price movements more often than chance would predict. The highest coherence is observed at the 2-day horizon across all windows (57,3%-61,4%). Coherence declines progressively from Day 5 onwards. Source: Authors

Horizon	7-day window	5-day window	2-day window
Day 1	35,8%	32,3%	35,2%
Day 2	60,3%	61,4%	57,3%
Day 3	48,2%	62,4%	50,7%
Day 4	61,5%	60,0%	58,3%
Day 5	54,8%	50,6%	52,7%
Day 6	48,8%	50,6%	50,0%
Day 7	47,5%	52,0%	48,1%
Day 8	41,4%	51,6%	42,9%

The best results are obtained with a 2-day horizon, with the coherence rate ranging from 57,3% to 61,4% across all windows. From the 4th day onward, a general downward trend is observed, particularly in the 7-day and 2-day windows. By the 8th day, coherence drops to 41,4% in the 7-day window and to 42,9% in the 2-day window.

This decline suggests that predicting price magnitude is inherently more challenging than predicting price direction. Media sentiment alone cannot capture how strongly markets will react, nor how fundamental factors such as earnings data or macroeconomic reports will amplify or suppress the signal. Consequently, sentiment-based interval forecasting requires additional inputs to improve accuracy over longer horizons (Abboud, 2025). Nevertheless, the results confirm that the prediction bands remain meaningful for horizons of up to four days.

5. Conclusion

This study aimed to examine whether aggregated media sentiment contains a measurable information signal regarding subsequent price movements in financial markets and under what conditions such signals are most reliable. The results, obtained through structured backtesting of 10,725 forecasts across three news-collection time windows, confirm the existence of consistent relational patterns, although their strength varies significantly by forecast horizon and signal type.

All three news collection time windows of 2, 5 and 7 days exceeded the random baseline of 50%, with the 7-day window achieving the highest overall directional accuracy of 61,2%. This finding supports the knowledge management perspective adopted in this study. Markets do not react to information immediately; rather, they respond when sufficient informational pressure from credible sources accumulates over time. The maximum accuracy of 67,1% was recorded for a 2-day forecasting horizon using a 7-day news window, further suggesting that the informational value of aggregated sentiment is most concentrated in the immediate period following its accumulation.

An analysis of signal type revealed significant asymmetry. The "Neutral" signal consistently demonstrated high accuracy across all time windows, ranging from 80,9% to 82,8%, suggesting that this framework is particularly effective at identifying states of information equilibrium. The weaker performance of the "Buy" and "Sell" signals, ranging from 27,6% to 43,5%, suggests that while the system reliably detects the absence of dominant directional pressure, distinguishing the directional pressure from informational noise remains a challenge. This finding is consistent with the broader sentiment analysis literature. It reinforces the view that sentiment-based systems are better understood as knowledge flow trackers rather than as deterministic forecasting tools.

Interval coherence analysis confirmed that prediction intervals are most reliable in the short term, up to four days, after which coherence gradually declines. This limitation reflects the inherent complexity of price forecasting, where fundamental factors such as earnings announcements and macroeconomic reports introduce variability that media sentiment alone cannot capture.

These findings suggest several avenues for future research. First, social media sentiment or earnings call transcripts could improve the accuracy of Buy and Sell signals. Second, applying this framework to markets outside the S&P 500 index, such as European or emerging markets, would test the generalizability of the findings. Third, combining sentiment signals with fundamental financial indicators may improve interval coherence over longer forecasting horizons.

Ethics declaration: This research did not require ethical clearance. The study relies exclusively on publicly available financial news data and stock price information. No human participants were involved in the research process.

AI declaration: Claude (Anthropic) was used in the development of this paper. The tool assisted with improving the language and clarity of the text and with structuring individual sections of the paper. All intellectual content, research design, data collection, analysis, and conclusions are the work of the authors.

References

Abboud, E. (2025) 'The Role of Artificial Intelligence in the Financial Industry: Enhancing Financial Analysis, Early Warning Detection, and Strategic Recommendations', SSRN Working Paper 5389089. https://ssrn.com/abstract=5389089

Ajayi, O.J. (2022) 'Knowledge Management Practices Among Stockbrokers', PhD thesis, Cardiff Metropolitan University. doi: https://doi.org/10.25401/cardiffmet.21799739

Araci, D. (2019) 'FinBERT: Financial Sentiment Analysis with Pre-trained Language Models', *arXiv preprint*, arXiv:1908.10063. https://arxiv.org/abs/1908.10063

Aroussi, R. (2024). yfinance: Download Market Data from Yahoo! Finance's API [software]. Available at: https://pypi.org/project/yfinance [Accessed 1 April 2026]

Buckman, S.R., Shapiro, A.H., Sudhof, M. and Wilson, D.J. (2020) 'News Sentiment in the Time of COVID-19', *FRBSF Economic Letter*, 2020-08. Federal Reserve Bank of San Francisco. https://www.frbsf.org/research-and-insights/publications/economic-letter/2020/04/news-sentiment-time-of-covid-19/

Chierici, R., Mazzucchelli, A., Garcia-Perez, A. and Vrontis, D. (2018) 'Transforming big data into knowledge: the role of knowledge management practice', *Management Decision*. doi:10.1108/MD-07-2018-0834.

Devlin, J., Chang, M-W., Lee, K. and Toutanova, K. (2019) 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding', *Proceedings of NAACL-HLT 2019*, pp. 4171–4186. doi: 10.18653/v1/N19-1423

Fama, E.F. (1970) 'Efficient Capital Markets: A Review of Theory and Empirical Work', *Journal of Finance*, 25(2), pp. 383–417. doi: 10.2307/2325486

Fama, E.F. (1991) 'Efficient Capital Markets: II', *Journal of Finance*, 46(5), pp. 1575–1617. doi: 10.2307/2328565

Fazlija, B. and Harder, P. (2022) 'Using Financial News Sentiment for Stock Price Direction Prediction', *Mathematics*, 10(13), p. 2156. doi:10.3390/math10132156.

Hutto, C.J. and Gilbert, E. (2014) 'VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text', *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media*. doi: 10.1609/icwsm.v8i1.14550

McKee, K. (2023). feedparser: Universal Feed Parser [software]. Available at: https://feedparser.readthedocs.io/en/latest/ [Accessed 1 April 2026]

Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L. and Trajanov, D. (2020) 'Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers', *IEEE Access*, 8, pp. 131662–131682. doi: 10.1109/ACCESS.2020.3009626

Ou-Yang, L. (2013) newspaper3k: Newspaper3k: Article scraping & curation [software]. Available at: https://newspaper.readthedocs.io/en/latest/ [Accessed 1 April 2026]

Puh, K. and Bagić Babac, M. (2023) 'Predicting stock market using natural language processing', *American Journal of Business*, 38(2), pp. 41–61. doi:10.1108/AJB-08-2022-0124.