

Leakage-Controlled Evaluation of Cross-Temporal Multimodal Fusion for Stock Movement Prediction

Yijun Lin*

Moansh University, Melbourne, Australia

**Corresponding author: Yijun Lin.*

Abstract

Financial news can complement price histories, but irregular timing makes multimodal stock-direction models vulnerable to leakage and overstatement. We evaluate Cross-Temporal Contrastive Alignment Fusion (CTCAF), which combines a five-day causal news memory with a recurrent price representation and a current-day gated cross-attention anchor. On public NIFTY data, three expanding walk-forward folds and five fixed seeds give descriptive Macro-F1 values of 0.467 for CTCAF, 0.495 for gated cross-attention fusion, and 0.510 for XGBoost; their seed-level standard deviations are 0.050, 0.038, and 0.021. The CTCAF-minus-gated fold-mean differences are minus 0.004, zero, and minus 0.080, so no advantage is observed at the independent-fold level. On StockNet-12, CTCAF and the gated baseline differ by only 0.001 Macro-F1. Ablations likewise show no consistent benefit from the contrastive term or longer text memory. Under this protocol, the added temporal-fusion components do not improve direction classification, underscoring the need for chronological controls and fold-level reporting in multimodal financial research.

Keywords

stock movement prediction, multimodal financial data, cross-temporal attention, gated fusion, walk-forward evaluation

1. Introduction

Next-day direction prediction tests whether market histories and public text contain complementary signals. Prices form regular numerical sequences, whereas news arrives asynchronously, may be absent on many days, and can refer to events whose effects unfold at different rates. Deep multimodal models therefore face two connected problems: matching information across time and separating a genuine contribution from capacity, threshold selection, or temporal leakage. Earlier systems combined word embeddings with price networks [1], contextualized financial language representations [2], and sentiment-derived features [3]. More recent work uses gated cross-attention to regulate the interaction between modalities [4]. These designs make temporal fusion plausible, but they do not guarantee that a longer text memory improves out-of-sample discrimination.

This study evaluates that proposition directly. We develop Cross-Temporal Contrastive Alignment Fusion (CTCAF), a model that conditions a five-day causal news memory on recurrent price tokens. Bidirectional

cross-attention supplies price-aware text and text-aware price representations; a learned age signal and forget gate control historical contributions; and an availability-aware multi-positive contrastive objective aligns observed text with price states without treating every other date as a negative. A frozen current-day gated cross-attention path remains available as an anchor. The construction is motivated by the possibility that relevant information is displaced across neighboring trading days, not by an assumption that the added machinery must outperform a simpler model.

The empirical question is consequently framed as a falsifiable comparison. On public NIFTY data, all preprocessing and model selection are restricted to each chronological training and validation segment. Three expanding walk-forward folds and five seeds expose variation across initialization and market period. Controls span majority class, logistic regression, XGBoost, unimodal neural models, early concatenation, unaligned dual attention, and gated cross-attention fusion (GCAF). We also evaluate the frozen pipeline on 12 alphabetically selected StockNet symbols. Macro-F1 is primary, with accuracy, Matthews correlation coefficient (MCC), ROC-AUC, balanced accuracy, and Brier score as complementary checks.

The comparison does not support the proposed improvement. CTCAF remains below GCAF and XGBoost on NIFTY, with a non-positive fold-mean difference from GCAF in all three folds. Removing the contrastive term changes little, while current-day text is numerically better overall. StockNet-12 shows only a small descriptive reversal. The contribution is therefore a leakage-controlled implementation and matched evaluation that identifies a protocol-specific limitation: neither longer history nor contrastive alignment supplies a consistent advantage here. The result provides neither a universal model ranking nor evidence of profitability.

2. Related Work

Financial-text prediction has progressed from aggregated sentiment and distributed word representations to contextual encoders and multimodal networks. Peng and Jiang [1] coupled news embeddings to deep price models, while Shah et al. [3] examined associations between news sentiment and market behavior. News–market interdependency also makes clear that text and price responses need not be simultaneous [5]. Contextual embeddings and label-aware training provide richer article representations [2, 6]; graph-based models add relations among firms and textual events [7]. Recent systems combine price and language-model signals directly [8, 9]. Public aligned resources now include the index-centered NIFTY headline collection [10] and StockNet’s company tweets and historical prices [11].

The fusion literature offers several ways to control cross-modal interactions. Low-rank transformer fusion reduces the cost of combining multimodal sequences [12], and gated cross-modal attention can suppress an unreliable channel while retaining conditional interactions [13]. For stock movement prediction, GCAF integrates price and text with a learned gate and is the closest controlling neural baseline [4]. Cross-temporal attention has also been explored for physiological sequences, where modalities may be displaced in time [14]. Cross-market models pursue a different source of context by transferring information among markets [15]. Together, these studies motivate temporal and gated fusion, but their distinct data and protocols make architecture-level performance expectations unsafe.

We isolate a narrower question: does cross-temporal news memory with contrastive alignment improve next-day direction classification over a matched current-day gate? The feature pipeline, chronology, seed set, selection metric, and external evaluation remain fixed. This constraint matters because an elaborate multimodal model can look favorable when it is compared only with weak baselines or evaluated under random splits. Cross-temporal fusion is therefore treated as a hypothesis, not a presumed source of advantage.

3. Method

3.1 Problem Formulation

For trading date t , let P_t denote a ten-day sequence of standardized market features and let $X_{t,d}$ denote news observed at lag $d \in \{0, \dots, 4\}$, where $d = 0$ is the current trading day. The binary target is

$$y_{t+1} = \mathbb{I}(c_{t+1} > c_t), \quad (1)$$

where c_t and c_{t+1} are consecutive released closing prices and $\mathbb{I}$ is the indicator function. Features and news available on date t predict the next released row, so weekends and exchange gaps are not synthesized. The model estimates $p_{t+1} = \Pr(y_{t+1} = 1 | P_t, X_{t,0:4})$. All normalizers, thresholds, checkpoints, and blending weights are selected without access to the test segment.

3.2 Causal Multimodal Encoding

A layer-normalized long short-term memory network maps P_t to price tokens H_t^p . Headlines or tweets assigned to one market date are concatenated into a daily document. The complete document is truncated to 128 tokens and encoded once with the frozen ProsusAI FinBERT checkpoint; its classification-token representation forms the daily vector. A trainable projection maps these vectors to $H_{t,d}^x$. Calendar age is represented by a learned lag embedding. Missing news receives a dedicated token and is masked from observed-text operations, so absence is represented consistently rather than replaced by future or neighboring content.

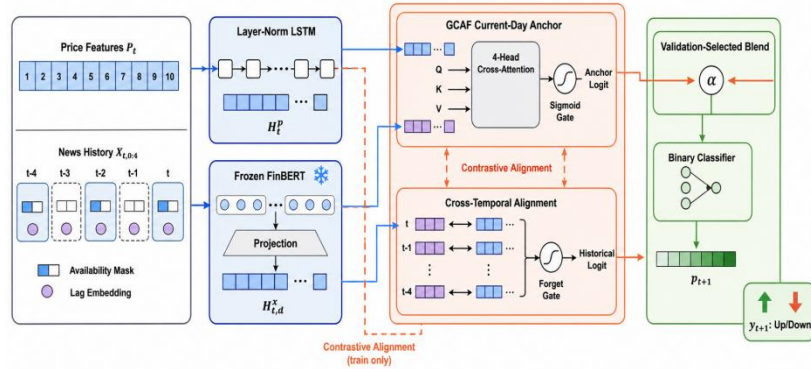
The current-day anchor follows GCAF. Projected news queries attend to price keys and values through four-head attention. A two-way sigmoid gate mixes the attended vector with pooled price information before binary classification. CTCAF retains this trained anchor and adds cross-temporal alignment. Text-to-price and price-to-text attention produce conditional representations for each observed lag. A forget gate then modulates the historical text contribution as

$$\tilde{h}_{t,d}^x = g_{t,d} \odot h_{t,d}^x, \quad g_{t,d} = \sigma(W_g[h_{t,d}^x; h_t^p; e_d] + b_g), \quad (2)$$

Here, $h_{t,d}^x$ is the price-aware text vector at lag d , h_t^p is the pooled text-aware price vector, and e_d is the lag embedding.

Figure 1: CTCAF architecture. Ten-day price features and a five-day causal news memory are encoded separately.

Bidirectional cross-attention, lag embeddings, and a forget gate form the historical branch; a current-day GCAF anchor provides a matched reference path. Only training examples with observed text enter the contrastive alignment objective



The matrix W_g and bias b_g are learned parameters. The symbol σ is the sigmoid function, $\odot$ denotes element-wise multiplication, and brackets denote concatenation. The gated history is pooled with observed-news masks and passed to a classifier. Figure 1 summarizes the causal data flow and the separation between the current-day anchor and historical alignment branch.

3.3 Availability-aware Alignment and Training

Identity-based contrastive learning creates false negatives when dates with the same movement label are forced apart. For anchor i , we instead define a positive set S_i containing other training examples with the same binary label and observed text. Examples without observed text are excluded from the contrastive denominator. Using normalized price and text projections z_i^p and z_j^x , the alignment loss is

$$\mathcal{L}_{\text{align}} = -\frac{1}{|\mathcal{A}|} \sum_{i \in \mathcal{A}} \log \frac{\sum_{j \in S_i} \exp((z_i^p)^\top z_j^x / \tau)}{\sum_{k \in \mathcal{A}, k \neq i} \exp((z_i^p)^\top z_k^x / \tau)}, \quad (3)$$

where $\mathcal{A}$ is the set of valid training anchors, k denotes the index of another valid anchor in the denominator, and $\tau = 0.1$ is the temperature. If an anchor has no valid positive, it does not contribute. The total objective is binary cross-entropy plus $0.1\mathcal{L}_{\text{align}}$. Labels and availability enter this objective only during training.

Both branches use a hidden dimension of 64 and dropout of 0.25. Optimization uses AdamW, a learning rate of 0.001, weight decay of 0.0001, batch size 64, and at most 24 epochs. Validation Macro-F1 selects checkpoints and decision thresholds; early stopping begins after epoch 20 with patience six. Thresholds are searched from 0.30 to 0.70 in increments of 0.005. A validation-selected scalar from $\{0, 0.25, 0.5, 0.75, 1\}$ blends the current-day anchor logit with the historical alignment logit. The historical branch can therefore be down-weighted without altering the controlling current-day implementation.

4. Experiments and Evidence

4.1 Data and Chronological Protocol

The public NIFTY collection [10] contains 2,111 raw daily rows and 2,101 usable next-day examples from 25 January 2010 to 17 September 2020; 55.12% of labels are positive. Its five leading raw price fields are open, high, low, close, and adjusted close. The context also contains volume, percentage change, MACD, upper and lower Bollinger bands, 30-day RSI, CCI and DX, and 30- and 60-day moving averages; the supplied row-level percentage change is appended as a sixteenth numerical input. The implementation applies $\log(1 + x)$ to the 60-day moving average before fold-specific standardization. The 2,101 daily text bundles contain a mean of 206.0 headlines (median 65; range 7–1,627), with no empty bundle. NIFTY provides trading dates but not intraday headline timestamps. We therefore use each date- t bundle only after that date's close to predict the next released trading row.

We use three expanding walk-forward folds. The first contains 1,050 training, 210 validation, and 210 test examples; its test dates run from 13 June 2016 to 29 June 2017. The second contains 1,260, 210, and 210 examples, with test dates from 5 July 2017 to 25 July 2018. The third contains 1,470, 210, and 421 examples, with test dates from 26 July 2018 to 17 September 2020. Training expands through the preceding validation period while every test segment remains later in calendar time. Preprocessing statistics are fitted within each training fold.

The external StockNet-12 set is derived from the public StockNet release [11] and uses the alphabetically first 12 eligible symbols: AAPL, ABB, ABBV, AEP, AGFS, AMGN, AMZN, BA, BABA, BAC, BBL, and BCH. It contains 5,522 examples, with 3,734 training, 600 validation, and 1,188 test examples. For each symbol-date, at most 24 unique tweets posted before 20:00 UTC are concatenated; 15,007 later messages are excluded. The symbol rule was fixed before outcome inspection.

Every neural result uses seeds 11, 29, 47, 71, and 101. Baselines comprise majority class, logistic regression, XGBoost, a price-only long short-term memory model, a news-only FinBERT classifier, early feature concatenation, unaligned dual attention, and GCAF. Means and standard deviations across the 15 fold-seed runs describe optimization variability, but the runs are not independent market replications. For each primary comparison, we first average the five seeds within each test fold and then report the three paired fold differences. With only three temporal folds, we make no significance or confidence-interval claim.

Feature construction is held constant across matched neural comparisons. Daily documents are encoded with the ProsusAI FinBERT checkpoint under Transformers 4.57.6 and cached by a content hash; all neural models receive the same current-day representation, while CTCAF can additionally query four earlier causal lags. Classical models use the flattened price window, 32 fixed text bins, an availability indicator, and log-transformed text count. Logistic regression uses $C = 0.5$, balanced class weights, and at most 1,000 iterations after a training-only scaler. XGBoost uses 160 trees, depth 3, learning rate 0.04, subsample and column fractions of 0.8, and $L_2 = 2$. These controls separate fusion structure from data access.

Validation Macro-F1 determines every neural threshold and checkpoint. The test split is evaluated once per frozen fold, model, and seed; it does not trigger refitting or hyperparameter selection. The completed record contains all 250 expected evaluations and their prediction archives, with finite values for every reported metric. This accounting includes the external and ablation conditions rather than only the primary comparison. It guards against selectively reporting seeds or variants after observing the outcome.

Metric roles are fixed before comparison. Macro-F1 gives equal weight to the two class-specific F1 scores; accuracy is retained to show the effect of the 55.12% positive majority. MCC and balanced accuracy summarize discrimination without relying on that majority rate, ROC-AUC evaluates score ranking, and Brier score measures probability error. No single secondary metric is used to reverse the primary decision. Fold-level contrasts preserve matching dates, while the within-fold seed spread is treated only as algorithmic variability.

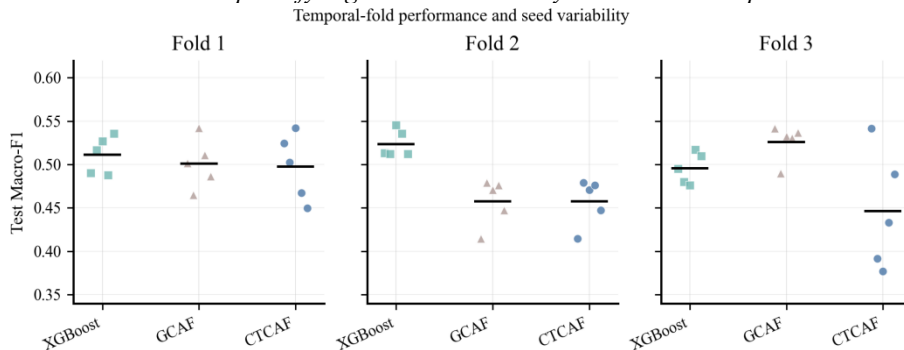
4.2 Primary Results

Table 1 summarizes the primary metrics. XGBoost has the highest Macro-F1 at 0.510 ± 0.021 ; GCAF is the strongest neural fusion baseline at 0.495 ± 0.038 . CTCAF attains 0.467 ± 0.050 , close to the price-only model and below both controls. Figure 2 separates the three test periods and displays all five seed runs within each fold. CTCAF’s MCC and ROC-AUC are also near their chance-centered values. The result does not support a gain from cross-temporal alignment.

Table 1: NIFTY walk-forward performance. Values are descriptive mean $\pm$ standard deviation across three folds and five seeds. Macro-F1 is primary; † marks the strongest neural method on that metric

Method	Accuracy	Macro-F1	MCC	ROC-AUC
Majority class	0.550 $\pm$ 0.037	0.354 $\pm$ 0.016	0.000 $\pm$ 0.000	0.500 $\pm$ 0.000
Logistic regression	0.504 $\pm$ 0.028	0.480 $\pm$ 0.039	0.051 $\pm$ 0.073	0.508 $\pm$ 0.027
XGBoost	0.525 $\pm$ 0.016	0.510 $\pm$ 0.021	0.042 $\pm$ 0.051	0.533 $\pm$ 0.017
Price LSTM	0.499 $\pm$ 0.030	0.467 $\pm$ 0.033	0.010 $\pm$ 0.030	0.503 $\pm$ 0.022
News FinBERT	0.492 $\pm$ 0.054	0.479 $\pm$ 0.044	−0.027 $\pm$ 0.087	0.486 $\pm$ 0.048
Early concatenation	0.513 $\pm$ 0.044	0.486 $\pm$ 0.032	0.005 $\pm$ 0.063	0.502 $\pm$ 0.044
Unaligned dual attention	0.488 $\pm$ 0.036	0.468 $\pm$ 0.043	−0.019 $\pm$ 0.065	0.487 $\pm$ 0.031
GCAF†	0.512 $\pm$ 0.048	0.495 $\pm$ 0.038	0.017 $\pm$ 0.071	0.516 $\pm$ 0.039
CTCAF	0.485 $\pm$ 0.041	0.467 $\pm$ 0.050	−0.008 $\pm$ 0.059	0.495 $\pm$ 0.032

Figure 2: NIFTY Macro-F1 by temporal fold. Each point is one of five optimization seeds; the black segment is the within-fold seed mean. Seeds quantify algorithmic variability and are not independent market samples

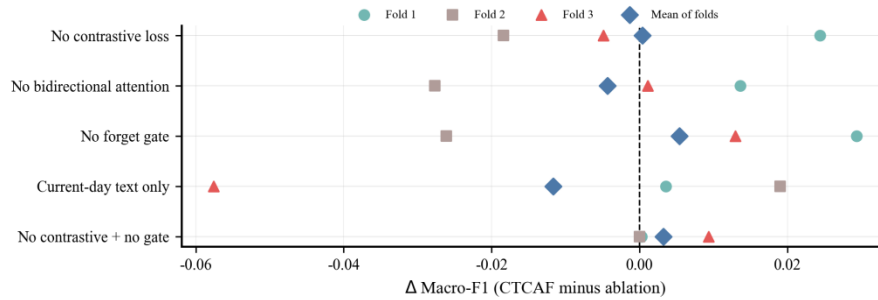


After averaging the five seeds within each test period, the CTCAF-minus-GCAF Macro-F1 differences across folds 1–3 are $(-0.004, 0.000, -0.080)$. The corresponding CTCAF-minus-XGBoost differences are $(-0.014, -0.066, -0.049)$. Thus, no temporal fold supports the planned improvement over either control. Three folds are insufficient for a stable sampling distribution, so these differences are reported without a p-value or confidence interval.

4.3 Ablation and External Evidence

Figure 3 reports fold-level effects of the proposed components. The mean-of-fold differences between CTCAF and the no-contrastive, no-bidirectional-attention, no-forget-gate, current-day-only, and joint no-contrastive/no-gate variants are $(0.000, -0.004, 0.005, -0.012, 0.003)$, respectively. No component effect is positive across all three folds. The current-day-only variant reaches 0.479 ± 0.046 overall and exceeds CTCAF in the latest test period, where their fold means differ by -0.058 .

Figure 3: Fold-level ablation effects. Points are differences between the CTCAF seed mean and the matched ablation seed mean within each temporal fold; diamonds are means of the three fold differences. No confidence interval is inferred from the three folds



The initially tested identity-based contrastive objective reaches 0.457878, whereas the availability-aware multi-positive version reaches 0.467103. This change avoids a known objective mismatch, but it does not turn the central result positive.

On StockNet-12, CTCAF and GCAF reach 0.524 ± 0.006 and 0.523 ± 0.004 Macro-F1, respectively. XGBoost, price-only LSTM, news-only FinBERT, and early concatenation obtain $(0.504 \pm 0.010, 0.510 \pm 0.008, 0.486 \pm 0.006, 0.513 \pm 0.009)$, in the same order. The CTCAF–GCAF gap is about 0.001 and is descriptive. Because the external evaluation uses one fixed symbol subset, it cannot overturn the primary NIFTY finding.

4.4 Robustness and Sensitivity

Post-training perturbations provide a further stress test without changing the fitted models. Setting all text embeddings and availability masks to zero gives an aggregated Macro-F1 of 0.362942 ± 0.075351 across the recorded NIFTY and StockNet-12 perturbation evaluations. Setting standardized columns after the five raw price fields to their training-fold mean, thereby removing derived indicators and the contemporaneous return feature, gives 0.440903 ± 0.081668 . These values are not retrained missing-modality baselines; they measure how frozen CTCAF predictions respond when one information source is removed at test time.

Fixed thresholds of 0.40, 0.45, 0.50, 0.55, and 0.60 yield aggregated Macro-F1 values of 0.375845, 0.399802, 0.445981, 0.400309, and 0.354555, respectively. The middle threshold is numerically best within this grid, but still does not change the primary comparison, which uses validation-selected thresholds separately within each fold. Both perturbations expose sensitivity to text availability and decision threshold. Neither supplies evidence that cross-temporal alignment itself is responsible for useful generalization.

4.5 Data and Code Availability

The NIFTY and StockNet source datasets are publicly available through the repositories cited above. Run-specific preprocessing and model code, configuration snapshots, chronological split indices, data fingerprints, metric tables, and prediction archives are retained with the study artifacts. No proprietary market feed is used.

5. Discussion

The experiments identify a protocol-specific limitation of multimodal financial fusion. A causal text history is attractive when news and prices react on different schedules, and related work supports both temporal context and gated interaction [4, 14]. Here, however, the five-day memory may add stale or weakly relevant text along with useful context. The current-day-only ablation is numerically stronger overall, and its fold differences change direction across periods. The availability-aware contrastive objective removes the false-negative pressure of identity matching, yet its removal leaves performance essentially unchanged. These patterns are consistent with limited incremental alignment signal; they do not demonstrate a causal mechanism.

The evaluation design supports that interpretation. Chronological folds prevent later market regimes from entering training, training-only preprocessing removes global normalization leakage, and fixed baselines separate architectural complexity from information access. MCC, ROC-AUC, and balanced accuracy also

show that the CTCAF result is not an isolated threshold artifact. All 250 planned evaluations completed with finite metrics and archived predictions.

The classical controls sharpen the result. XGBoost leads the primary Macro-F1 comparison despite lacking learned cross-modal attention, while logistic regression remains competitive with several neural variants. This does not imply that tree ensembles are generally preferable for financial prediction. It shows that the available sample and feature construction do not reward the added CTCAF capacity under the selected chronology. Likewise, StockNet-12 cautions against reading one dataset as a universal verdict: its ordering changes slightly, but the gap is too small and the replication structure too narrow to support a stronger conclusion.

The scope remains limited. NIFTY represents one index-centered historical period, and StockNet-12 is a fixed subset rather than a broad cross-market replication. Seeds within a fold share dates, leaving only three primary temporal units for comparison. FinBERT is frozen, each daily document is truncated to 128 tokens, and the five-day memory uses one cutoff policy. NIFTY contains no empty daily news bundle, so it cannot test natural missing-news behavior. The study does not estimate transaction costs, returns, risk-adjusted performance, or deployability; classification metrics cannot support a profitability claim.

6. Conclusion

CTCAF aligns asynchronous financial text with price histories through bidirectional cross-attention, lag-aware forgetting, and availability-aware contrastive learning. Under the leakage-controlled NIFTY protocol, it reaches 0.467 ± 0.050 Macro-F1 and remains below both GCAF and XGBoost. Fold-level comparisons and ablations provide no consistent benefit from the contrastive term or longer text memory. A 0.001 StockNet-12 difference is descriptive and does not change the primary conclusion. For these datasets, dates, and preprocessing choices, added cross-temporal complexity does not improve movement classification. Broader temporal and cross-market evidence would be required for a stronger claim.

References

- [1] Peng, Y., and Jiang, H., “Leverage financial news to predict stock price movements using word embeddings and deep neural networks,” *arXiv preprint arXiv:1506.07220*, 2015.
- [2] Chen, Q., “Stock movement prediction with financial news using contextualized embedding from BERT,” *arXiv preprint arXiv:2107.08721*, 2021.
- [3] Shah, D., Isah, H., and Zulkernine, F., “Predicting the effects of news sentiments on the stock market,” *arXiv preprint arXiv:1812.04199*, 2018.
- [4] Zong, C., Wan, J., Cascone, L., and Zhou, H., “Stock movement prediction with multimodal stable fusion via gated cross-attention mechanism,” *Complex & Intelligent Systems*, 11, 396, 2025, <https://doi.org/10.1007/s40747-025-02023-3>.
- [5] Hwang, E., and Kim, Y.-H., “Interdependency between the stock market and financial news,” *arXiv preprint arXiv:1909.00344*, 2019.
- [6] Villamil, L., Bausback, R., Salman, S., Liu, T. L., Horn, C., and Liu, X., “Improved stock price movement classification using news articles based on embeddings and label smoothing,” *arXiv preprint arXiv:2301.10458*, 2023.
- [7] Chen, Q., and Robert, C.-Y., “Graph-based learning for stock movement prediction with textual and relational data,” *The Journal of Financial Data Science*, 4, (4), 152–166, 2022, <https://doi.org/10.3905/jfds.2022.1.104>.
- [8] Chen, Q., “Stock price prediction with LLM-guided market movement signals and transformer model,” *FinTech and Sustainable Innovation*, 1, A12, 2025, <https://doi.org/10.47852/bonviewFSI52025703>.
- [9] Elahi, A., and Taghvaei, F., “Combining financial data and news articles for stock price movement prediction using large language models,” *arXiv preprint arXiv:2411.01368*, 2024.

- [10] Saqur, R., Kato, K., Vinden, N., and Rudzicz, F., “NIFTY financial news headlines dataset,” *arXiv preprint arXiv:2405.09747*, 2024, <https://arxiv.org/abs/2405.09747>.
- [11] Xu, Y., and Cohen, S. B., *Stock movement prediction from tweets and historical prices*, *Proceedings of the 56th annual meeting of the association for computational linguistics*, 1970–1979, Association for Computational Linguistics, 2018, <https://doi.org/10.18653/v1/P18-1183>.
- [12] Sahay, S., Okur, E., Kumar, S. H., and Nachman, L., “Low rank fusion based transformers for multimodal sequences,” *arXiv preprint arXiv:2007.02038*, 2020.
- [13] Jiang, M., and Ji, S., “Cross-modality gated attention fusion for multimodal sentiment analysis,” *arXiv preprint arXiv:2208.11893*, 2022.
- [14] Khorasani, A., and Demazure, T., “Cross-temporal attention fusion (CTAF) for multimodal physiological signals in self-supervised learning,” *arXiv preprint arXiv:2602.02784*, 2026.
- [15] Lin, S., Chen, Y., Qi, Y., Ma, C., Cao, B., Zhang, Y., Liu, X., and Guo, J., “CSPO: Cross-market synergistic stock price movement forecasting with pseudo-volatility optimization,” *arXiv preprint arXiv:2503.22740*, 2025.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

