

Sentiment-Weighted Fusion of Financial News and Technical Indicators for Short-Horizon Equity Trend Classification: An Evidence-Informed Research Framework

Poornima Chourasia^{1*}, Yogendra Singh Rajavat²

¹Institute of Computer Science, Samrat Vikramaditya University, Ujjain, Madhya Pradesh, India; Assistant Professor, Prashanti College of Professional Studies, Ujjain, India.

Email: poornimachourasiasharma@gmail.com; ORCID: 0009-0006-6172-0935

²Prestige Institute of Management & Research, Dewas, Madhya Pradesh, India; Research Guide, Samrat Vikramaditya University, Ujjain, Madhya Pradesh, India. ORCID: 0009-0001-8481-9261

Abstract: Short-horizon equity trend prediction remains difficult because price formation reflects both historical market behaviour and rapidly changing information flows. Technical indicators condense past open-high-low-close-volume data into interpretable measures of trend, momentum, volatility and participation. At the same time, financial news sentiment captures the market relevance of textual disclosures, macroeconomic updates and company-specific events. This paper develops a sentiment-weighted fusion framework that combines transformer-derived financial news sentiment with selected technical indicators for next-day equity trend classification. The proposed design uses a finance-domain transformer, such as FinBERT, to transform news headlines or short articles into positive, negative and neutral sentiment probabilities; these outputs are then aggregated by trading date and merged with engineered indicators including moving averages, relative strength index, moving average convergence divergence, Bollinger Bands, daily return, volume change and historical volatility. Instead of presenting unverified or simulated accuracy values, the manuscript positions the contribution as an evidence-informed methodological framework supported by established studies on public mood, LSTM-based market forecasting, transformer-based sentiment classification and hybrid sentiment-technical prediction. The paper specifies the research gap, hypotheses, input variables, feature construction, validation design, evaluation metrics and reproducibility requirements for future empirical implementation. The framework is intended to support a narrower publishable subset of a broader multimodal financial market prediction research agenda while maintaining transparency about the need for original dataset validation before final performance claims are made.

Keywords: financial news sentiment; FinBERT; technical indicators; stock trend classification; transformer model; feature fusion; financial forecasting; machine learning

1. Introduction

Equity market prediction is a long-standing research problem in computational finance, investment analytics and risk management. Although historical price and volume data are widely used for trend identification, short-term price movement is also shaped by new information, including earnings announcements, regulatory updates, monetary policy signals, geopolitical developments, analyst commentary and sector-specific news. A prediction model that depends only on numerical market data may therefore miss an important part of the decision environment in which investors operate.

Technical analysis converts raw market observations into structured signals. Moving averages indicate trend direction, the relative strength index captures momentum, moving average convergence divergence reflects the interaction of short-term and long-term trends, Bollinger Bands represent volatility around a moving average, and

volume-based measures describe market participation. These indicators are useful because they are interpretable and easily aligned with daily trading data. However, they are backward-looking by construction and may react slowly to abrupt news-driven market shifts.

Financial news sentiment analysis offers a complementary view. Modern natural language processing methods can convert unstructured news into numerical sentiment features that are suitable for machine learning. Finance-specific transformer models are particularly relevant because financial text is context sensitive: words that appear negative in ordinary language can have different meaning in corporate filings, credit reports or market commentary. By learning contextual word representations, transformer models can improve the quality of sentiment extraction compared with general dictionaries or simple polarity counts.

The present paper narrows the broader research area of multimodal financial market prediction into a focused manuscript on sentiment-weighted technical indicator fusion for short-horizon equity trend classification. The study is designed as a transparent framework paper rather than a simulated experiment. It avoids fabricated performance values and instead identifies a reproducible pathway for empirical testing. This positioning is important for journal submission because performance claims in financial forecasting must be supported by clearly documented data sources, train-test splits, implementation details and robustness checks.

The objectives of this study are to: (i) formulate a publishable feature-fusion framework that combines financial news sentiment and technical indicators; (ii) specify a transformer-based sentiment extraction procedure suitable for financial text; (iii) define technical variables and a next-day trend classification target; (iv) provide hypotheses and evaluation criteria for future empirical validation; and (v) synthesize prior evidence that justifies the proposed design without overstating unsupported results.

2. Literature Review

2.1 Technical indicators and machine-learning models

Machine-learning methods have become common in stock market research because financial time series often contain non-linear relationships, time dependence and regime changes. Classical classifiers such as logistic regression provide interpretable baselines, whereas support vector machines, random forests, gradient boosting and extreme gradient boosting can model complex feature interactions. Recurrent neural networks, especially long short-term memory models, are frequently used when the temporal order of observations is central to the forecasting problem.

Fischer and Krauss (2018) demonstrated the relevance of LSTM networks for directional prediction of S&P 500 constituent stocks and reported stronger results than several memory-free classifiers in their setting. Their work supports the broader argument that sequential modelling can capture patterns that are not fully represented in one-day lagged features. At the same time, the financial forecasting literature consistently warns that predictive performance depends heavily on the market, period, costs, data-cleaning choices and validation design.

2.2 Financial news sentiment and transformer models

Financial news is an important information channel because it shapes investor beliefs about expected cash flows, risk and uncertainty. News about earnings, management changes, legal disputes, interest rates, inflation, mergers, credit quality or sector policy can alter short-term demand for a stock even when recent technical signals appear stable. Sentiment analysis attempts to quantify these qualitative signals by assigning positive, negative or neutral scores to textual observations.

Earlier sentiment approaches relied on general-purpose dictionaries or finance-specific lexicons. Although these methods are transparent, they can struggle with negation, domain-specific terminology and context-dependent expressions. Transformer-based models address these limitations by using attention mechanisms and contextual representations. FinBERT, introduced by Araci (2019), adapts BERT-style language modelling to financial sentiment analysis and provides a practical foundation for extracting sentiment probabilities from market-related text.

2.3 Multimodal and hybrid market prediction

A growing strand of research combines textual information with numerical market data. The rationale is that price-volume indicators reflect historical behaviour, while sentiment features represent newly released information and investor interpretation. Hybrid designs may use early fusion, where features are merged before model training; late fusion, where separate model outputs are combined; or attention-based fusion, where the model learns how much importance to assign to each modality over time.

Bollen, Mao and Zeng (2011) showed that public mood signals can be informative for market direction. Hu et al. (2018) used a news-oriented deep learning architecture for stock trend prediction, and Kabbani and Usta (2022) directly combined news sentiment with technical indicators in a machine-learning environment. These studies support the idea that sentiment and technical variables should not be treated as isolated research streams. However, reported results are often data-source-specific, market-specific and period-specific, which creates a need for careful reproducibility and restrained interpretation.

3. Research Gap and Contributions

The literature supports the usefulness of technical indicators, financial sentiment and deep learning, but four gaps remain relevant. First, many studies investigate text and numerical market data separately, limiting the ability to model their interaction. Second, some hybrid studies use simple polarity scores rather than finance-domain transformer outputs. Third, many published experiments are difficult to compare because they use different markets, news sources, time windows, target definitions and validation strategies. Fourth, manuscripts that are derived from broad doctoral topics often need a narrower, journal-ready scope that does not duplicate the dissertation title or make unsupported empirical claims.

This paper addresses these gaps by proposing a sentiment-weighted technical indicator fusion framework with a clear data pipeline, feature definitions, hypotheses and validation plan. The contribution is methodological and evidence-informed: it creates a precise structure that can be implemented using real financial news and market datasets, while distinguishing between literature-supported justification and original empirical proof.

4. Proposed Framework

The proposed framework contains two parallel feature-generation branches followed by temporal alignment, early fusion and trend classification. The first branch converts financial news into daily sentiment features, while the second branch generates technical indicators from OHLCV market data. The fused feature matrix is then used to train and evaluate classifiers for next-day directional movement.

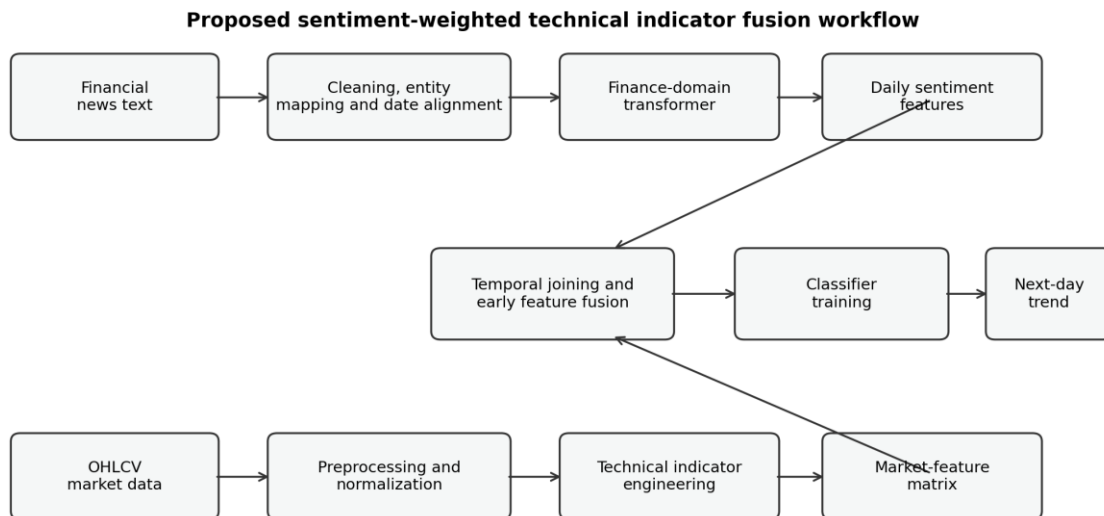


Figure 1. Proposed sentiment-weighted technical indicator fusion workflow.

4.1 Textual sentiment branch

The textual branch accepts financial news headlines, summaries or short articles that can be mapped to a company, sector or index. Preprocessing includes duplicate removal, source verification, time-stamp normalization, removal of irrelevant symbols, and mapping of each item to the correct trading date. Weekend and holiday news may

be assigned to the next trading session when the research design focuses on market reaction rather than publication timing alone.

A finance-domain transformer model is then applied to each news item. The output can be represented as three probabilities: P(positive), P(negative) and P(neutral). A compact sentiment score can be computed as P(positive) minus P(negative), while the original probability vector may also be retained to preserve more information. When multiple news items are available for the same instrument and day, the model can aggregate them using a simple average, source-weighted average or time-decay weighting.

4.2 Numerical market branch

The numerical branch uses open, high, low, close, adjusted close and volume data. After removing non-trading days and handling missing values, technical indicators are generated to represent trend, momentum, volatility, return and participation. The indicators selected for the baseline implementation are intentionally standard so that the framework remains reproducible and interpretable.

The proposed indicators include short-term and medium-term moving averages, RSI, MACD, Bollinger Bands, daily return, volume change and rolling volatility. These features can be normalized using training-set statistics to avoid information leakage. When sequential models are used, the input matrix may be converted into rolling windows so that the model receives a short history of fused observations rather than a single-day record.

4.3 Fusion and classification layer

The daily sentiment features and technical features are joined by instrument and trading date. Early feature fusion is recommended for the first implementation because it is transparent, easy to reproduce and compatible with both classical classifiers and neural models. The fused vector for day t can be represented as: $X(t) = [\text{Technical indicators}(t), \text{Sentiment probabilities}(t), \text{Aggregated sentiment score}(t)]$.

The target variable is next-day direction. If $\text{Close}(t+1)$ is greater than $\text{Close}(t)$, the observation is labelled as Uptrend; otherwise, it is labelled as Downtrend. This binary formulation is suitable for short-horizon classification and allows comparisons across technical-only, sentiment-only and fused feature sets.

5. Methodology for Empirical Implementation

The framework can be implemented as an empirical study using selected stocks or indices from NIFTY 50, NIFTY Bank, NIFTY IT, S&P 500, NASDAQ or another defined universe. The final selection should be based on data quality, market relevance, news availability and the target journal scope. The recommended workflow is presented in Table 1.

Table 1. Recommended empirical workflow for implementing the proposed framework

Step	Procedure	Expected output
1	Define market universe, study period, data sources and inclusion criteria.	Replicable research boundary.
2	Collect financial news with source, publication date, entity tag and headline/body text.	Raw textual dataset.
3	Collect OHLCV market data for the same universe and time period.	Raw numerical dataset.
4	Clean, deduplicate and align news with trading days.	Date-aligned textual observations.
5	Extract transformer-based sentiment probabilities and daily sentiment aggregates.	Sentiment feature table.
6	Generate technical indicators and normalize features using training data only.	Numerical feature table.

7	Create next-day directional labels and split data chronologically.	Training, validation and test sets.
8	Train technical-only, sentiment-only and fused models.	Comparable prediction outputs.
9	Evaluate classification metrics, robustness and, if relevant, trading-cost-adjusted performance.	Evidence for hypotheses and model usefulness.

Table 2 defines the main variables. The exact window lengths for indicators should be chosen before model training and reported transparently to avoid post-hoc selection bias.

Table 2. Proposed feature groups and modelling role

Feature group	Examples	Purpose in the framework
Trend	Moving average, moving-average gap, price relative to moving average	Captures short-term and medium-term direction.
Momentum	RSI, MACD, MACD signal difference	Represents speed and strength of recent price movement.
Volatility	Bollinger Band width, rolling standard deviation of return	Measures market uncertainty and dispersion.
Return and volume	Daily return, lagged return, volume change, rolling volume average	Captures immediate movement and market participation.
Sentiment probability	Positive, negative and neutral transformer probabilities	Preserves model confidence for each sentiment class.
Aggregated sentiment	Daily average score, time-weighted score, negative-news intensity	Converts multiple news items into trading-date features.
Target	Uptrend or Downtrend based on $Close(t+1)$ relative to $Close(t)$	Defines the supervised classification problem.

5.1 Hypotheses

H1: Transformer-derived financial news sentiment contains information relevant to short-horizon equity trend classification.

H2: Technical indicators derived from OHLCV data provide useful numerical signals for next-day directional movement.

H3: A fused sentiment-technical feature set provides a stronger prediction foundation than either modality used alone.

H4: Negative sentiment is expected to show stronger short-term association with price declines during volatile periods than positive sentiment shows with price increases.

H5: Final performance claims for the proposed framework require original empirical validation with documented data, code and evaluation settings.

5.2 Model comparison and evaluation

A journal-ready empirical implementation should compare at least three feature settings: technical-only features, sentiment-only features and fused sentiment-technical features. Suitable baseline models include logistic regression and support vector machine. Tree-based models such as random forest and XGBoost can capture non-linear interactions, while LSTM or temporal fusion transformer architectures can be considered when rolling-window sequences are used.

The minimum classification metrics should include accuracy, precision, recall, F1-score and ROC-AUC. Because class imbalance is common in directional market data, F1-score and balanced accuracy should be reported in

addition to raw accuracy. If the study evaluates trading usefulness, it should also report transaction-cost-adjusted return, maximum drawdown and Sharpe ratio. Chronological train-test splitting is recommended because random splitting can leak future market information into the training set.

6. Evidence Synthesis from Prior Studies

This manuscript does not report simulated performance values. Instead, Table 3 summarizes evidence from established research that supports the feasibility of the proposed fusion design. The table should be interpreted as justification for the framework, not as proof of the performance of the present model.

Table 3. Evidence base supporting the sentiment-technical fusion approach

Study	Main data or method	Key reported contribution	Relevance to this framework
Bollen, Mao and Zeng (2011)	Public mood signals and DJIA movement	Reported that mood indicators improved directional market prediction.	Supports sentiment as a market-relevant information signal.
Loughran and McDonald (2011)	Finance-specific textual analysis and 10-K terminology	Showed that finance-domain language differs from general dictionaries.	Supports the need for domain-aware sentiment treatment.
Fischer and Krauss (2018)	LSTM models for S&P 500 constituent stocks	Demonstrated the value of sequential learning for financial direction prediction.	Supports time-series modelling in the numerical branch.
Araci (2019)	FinBERT for financial sentiment classification	Introduced a finance-domain transformer approach for sentiment analysis.	Supports transformer-based extraction in the textual branch.
Hu et al. (2018)	News-oriented stock trend prediction with attention mechanisms	Used deep learning to incorporate sequential news information.	Supports the role of news and attention in market prediction.
Lim et al. (2021)	Temporal fusion transformer for multi-horizon forecasting	Proposed an interpretable attention-based time-series architecture.	Supports advanced temporal fusion extensions.
Kabbani and Usta (2022)	Technical indicators and daily news sentiment in Spark	Reported that a hybrid feature set can be useful for stock trend prediction.	Directly supports combining sentiment and technical indicators.
Yang, Liu and Wang (2023)	Open-source financial large language models	Highlighted the growing use of domain-focused financial language models.	Supports future extensions using financial LLMs and explainable sentiment features.

7. Discussion

The proposed framework is academically relevant because it combines two complementary views of market behaviour. Technical indicators summarize what has already occurred in price and volume, while financial news sentiment represents newly released information and investor interpretation. A fused design can therefore provide a richer representation of the information environment surrounding short-term price movement.

The framework is also practically useful because its components are interpretable. Traders and analysts are familiar with moving averages, RSI, MACD, Bollinger Bands and volume measures. Sentiment probabilities are

likewise understandable when presented as daily positive, negative and neutral signals. This interpretability is important in financial applications where users often require explanations rather than black-box predictions alone.

For Scopus-indexed journal submission, the strongest route is to implement the framework empirically and present original results. If the target journal accepts conceptual or methodological papers, the present manuscript can be submitted as an evidence-informed framework paper. If the journal expects experimental evidence, the current evidence-synthesis section should be replaced or supplemented with original dataset results, statistical tests and reproducibility materials.

8. Limitations and Future Research

The main limitation is that this manuscript does not claim original model performance. It deliberately avoids simulated accuracy, invented confusion matrices or unsupported trading returns. This improves research integrity but means that the framework requires empirical implementation before it can be evaluated as a predictive system.

A second limitation concerns data quality. Financial news sources differ in speed, reliability, coverage and relevance. Incorrect entity mapping or delayed publication timestamps can distort sentiment alignment. Third, transformer models may still misclassify financial nuance, sarcasm, forward-looking uncertainty or mixed-sentiment statements. Fourth, market movement is also influenced by macroeconomic variables, institutional flows, liquidity conditions and geopolitical events that may not be captured by news sentiment and technical indicators alone.

Future research should test the framework across multiple market segments and time periods, use walk-forward validation, evaluate transaction costs and slippage, and compare early fusion with late fusion and attention-based fusion. Explainable artificial intelligence methods such as SHAP or LIME can be used to identify which sentiment and technical variables contribute most to the final prediction. Extensions may also include analyst reports, social media sentiment, financial ratios, volatility indices and macroeconomic indicators.

9. Conclusion

This paper presented a sentiment-weighted fusion framework for short-horizon equity trend classification. The framework integrates finance-domain transformer sentiment features with technical indicators derived from OHLCV market data and defines a reproducible pathway for next-day directional prediction. It is intentionally narrower than a full multimodal doctoral research topic and is therefore suitable as a focused journal manuscript.

The paper contributes by clarifying the research gap, feature pipeline, hypotheses, target definition, evaluation design and evidence base for sentiment-technical fusion. It does not report unverified experimental accuracy. Instead, it provides an ethically framed and methodologically transparent foundation for future implementation. Final journal submission should include original data collection, code environment details, model outputs, statistical evaluation and journal-specific formatting according to the selected Scopus-indexed journal.

Declarations

Conflict of interest: The authors declare that they have no known competing financial or personal interests that could have influenced the work reported in this manuscript.

Funding: No external funding is reported for this manuscript draft.

Data availability: This manuscript presents an evidence-informed research framework. Original datasets, code and processed features should be supplied or described when the framework is implemented empirically.

Ethical statement: The manuscript is based on publicly available literature and proposes a computational framework. No human-subject data is collected in this draft.

Author contributions: Prof. Poornima Chourasia contributed to conceptualization, manuscript preparation and research framing. Dr. Yogendra Singh Rajavat contributed to supervision, guidance and critical academic review.

References

1. Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models. arXiv preprint arXiv:1908.10063. <https://arxiv.org/abs/1908.10063>
2. Bollen, J., Mao, H., & Zeng, X. J. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8. <https://doi.org/10.1016/j.jocs.2010.12.007>

3. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*, 4171-4186.
4. Ding, X., Zhang, Y., Liu, T., & Duan, J. (2015). Deep learning for event-driven stock prediction. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2327-2333.
5. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654-669. <https://doi.org/10.1016/j.ejor.2017.11.054>
6. Hu, Z., Liu, W., Bian, J., Liu, X., & Liu, T.-Y. (2018). Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 261-269. <https://doi.org/10.1145/3159652.3159700>
7. Jiang, W. (2021). Applications of deep learning in stock market prediction: Recent progress. *Expert Systems with Applications*, 184, 115537. <https://doi.org/10.1016/j.eswa.2021.115537>
8. Kabbani, T., & Usta, F. E. (2022). Predicting the stock trend using news sentiment analysis and technical indicators in Spark. *arXiv preprint arXiv:2201.12283*. <https://arxiv.org/abs/2201.12283>
9. Lim, B., Arik, S. O., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748-1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
10. Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
11. Schumaker, R. P., & Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news. *ACM Transactions on Information Systems*, 27(2), Article 12. <https://doi.org/10.1145/1462198.1462204>
12. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
13. Xu, Y., & Cohen, S. B. (2018). Stock movement prediction from tweets and historical prices. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 1970-1979*. <https://doi.org/10.18653/v1/P18-1183>
14. Yang, H., Liu, X.-Y., & Wang, C. D. (2023). FinGPT: Open-source financial large language models. *Proceedings of the ACM International Conference on AI in Finance*. <https://arxiv.org/abs/2306.06031>