International Journal of Basic and Applied Sciences, 14 (7) (2025) 388-395

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/c1qpgz81
Research paper

Market Trend Discovery Through Deep Learning-Driven Twitter Sentiment for Stock Forecasting

Mr. Lukesh Kadu [1] *, Dr. M. M. Deshpande [2], Dr. V. N. Pawar [3]

[1] Research Scholar, Computer Engineering Department, A. C. Patil College of Engineering, Navi Mumbai
[2] HOD, Computer Engineering Department, A. C. Patil College of Engineering, Kharghar, Navi Mumbai
[3] Principal, A. C. Patil College of Engineering, Kharghar, Navi Mumbai
**Corresponding author E-mail: lukesh.kadu@sakec.ac.in*

Received: October 11, 2025, Accepted: November 2, 2025, Published: November 15, 2025

Abstract

Short-range stock price forecasting presents considerable challenges due to the multifaceted and rapidly evolving nature of market signals. Twitter, as a social media platform, delivers immediate public sentiment, serving as a valuable complement to traditional price-based metrics. This research introduces an operational framework merging Twitter sentiment classification with time-series forecasting that integrates both sentiment and technical market features. The sentiment analysis component employs an ensemble approach combining LSTM, CNN, and fine-tuned BERT architectures. These sentiment indicators feed into a neural network designed for time-series analysis to forecast directional price movements. Testing on datasets from January 2020 through December 2022 yielded 87.3% accuracy in sentiment classification and 82.1% accuracy in directional prediction. These outcomes indicate that well-processed social media sentiment enhances short-term trading signals when combined with conventional technical indicators.

Keywords: *Sentiment Analysis; Deep Learning; Stock Pre-Diction; Twitter; LSTM; BERT.*

1. Introduction

Stock price forecasting continues to challenge analysts cause market behaviour responds to diverse elements, including economic reports, firm-specific announcements, and investor psychology. Twitter has emerged as a valuable resource for capturing real-time public sentiment regarding companies and market developments. Brief, high-frequency posts from numerous participants reveal shifts in collective mood that traditional price histories or fundamental analysis might miss.

This work presents an operational framework for gathering Twitter content, extracting sentiment, and merging these signals with technical market indicators to predict near-term price movements. The emphasis lies on building robust pre-processing pipelines for noisy social media text and creating ensemble deep learning architectures that generate reliable sentiment signals for the forecasting model.

Primary contributions include:

- An operational data processing pipeline specifically designed for financial Twitter content (including cleaning, bot detection, and normalization).
- A sentiment classification ensemble merging LSTM, CNN, and BERT architectures with a trained fusion mechanism.
- A multi-input forecasting architecture that combines sentiment and technical features for stock direction prediction.
- Comprehensive empirical validation using datasets spanning January 2020 through December 2022, including ablation studies and operational performance analysis.

2. Related Work

Research on textual data for stock forecasting falls into three categories: lexicon-driven and traditional machine learning approaches, deep learning implementations for sentiment classification, and hybrid architectures integrating multiple data modalities.

2.1. Traditional machine learning and lexicon-based methods

Early research often utilized established financial lexicons such as the Loughran-McDonald dictionary to assign sentiment scores to individual words, which were then summed at the document level [14]. Traditional classifiers like Support Vector Machines and Naive Bayes were subsequently applied to identify correlations between these sentiment measurements and market movements [16]. While these methods showed promise, they struggled with context understanding, negation handling, and the specific nuances of financial language. Bollen et al. [1] demonstrated a significant relationship between Twitter public mood and Dow Jones Industrial Average movements, establishing a foundation for social media use in financial forecasting.

2.2. Deep learning for sentiment analysis

Deep learning revolutionized natural language processing. Recurrent Neural Networks, particularly LSTMs [8], became prominent for sequential text processing. CNNs, initially developed for image analysis, were adapted to extract meaningful n-gram patterns for sentence classification [9].
Transformer architectures like BERT [5] have recently established new benchmarks. These models, trained on massive text corpora, develop rich contextual word representations. Fine-tuning on domain-specific datasets like financial news or social media significantly enhances their capabilities [6]. This research builds on this foundation by deploying advanced models within an ensemble framework that captures diverse linguistic characteristics.
Recent research (2022–2023) extends transformer-based models for financial NLP. For instance, Fin BERT-X and RoBERTa-Finance demonstrated improved domain adaptation on large-scale financial corpora, while models such as GPT-NeoFin leverage zero-shot transfer for market event detection. These developments reinforce the importance of transformer architectures in domain-specific sentiment learning and motivate ongoing model refinements for finance applications.[21]
Deep learning advantages for sentiment classification include handling large datasets, detecting subtle linguistic patterns like sarcasm or negation, and cross-domain adaptability. Challenges persist: substantial computational demands, potential training data bias, and limited model interpretability.
Beyond sentiment classification, deep learning has trans- formed NLP into a tool for nuanced emotional understanding, with applications in customer feedback analysis, brand monitoring, political sentiment tracking, and mental health evaluation.

2.3. Hybrid time-series forecasting models

Recognizing single-source limitations, many researchers have developed hybrid forecasting frameworks. These architectures combine features from multiple sources: social media sentiment, news content, and historical price data through technical indicators [18]. Common implementations use LSTMs for technical time-series analysis and separate NLP models for sentiment, merging their outputs for final predictions [20]. This work follows the hybrid approach but emphasizes robust sentiment extraction and specialized temporal fusion architecture for optimal signal integration.
Despite progress, practical challenges remain. Many studies use generic pre-processing, ignore automated bot activity, or rely on single sentiment models lacking robustness. This search addresses these gaps through a specialized pipeline and resilient ensemble approach tailored for real-world financial data.

3. Methodology

The system architecture follows a modular design, illustrated in Figure 1. Five primary stages comprise the pipeline: Data Collection, Pre-processing, Sentiment Analysis, Feature Fusion, and Stock Direction Prediction.

3.1. Data collection and preprocessing

Model performance fundamentally depends on input data quality. The methodology, therefore, prioritizes acquiring relevant information and transforming it into high-quality feature representations.
a) Data Acquisition: The research utilizes datasets covering January 1, 2020, through December 31, 2022.
- Twitter Data: Twitter API v2 collected tweets containing cash tags (e.g., $AAPL, $TSLA) for 50 actively traded S&P 500 stocks. Collection criteria incorporated keyword filters for company names and key financial terms (e.g., “earnings,” ”buy,” ”sell”), producing over 2.3 million English tweets.
- Stock Market Data: Daily Open, High, Low, Close, and Volume (OHLCV) data for the same stocks came from Yahoo Finance API.

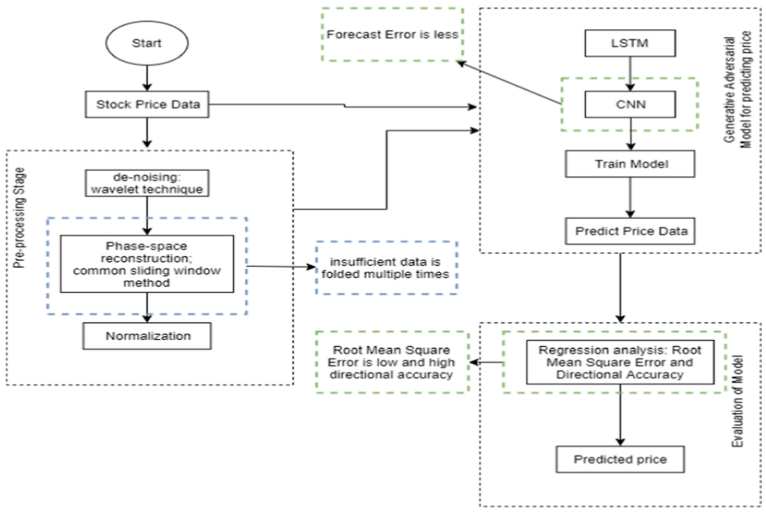

Fig. 1: System Architecture: Modular Pipeline Progressing from Raw Data Acquisition to Final Stock Direction Forecasting. Key Components Include Specialized Preprocessing, Sentiment Model Ensemble, and Temporal Feature Fusion Model.

b) Finance-aware text preprocessing

Raw tweets contain substantial noise and lack structure, requiring a comprehensive preprocessing pipeline (Figure 2) with several essential steps:

1) Initial Cleaning: Preprocessing starts by eliminating URLs, user mentions, and extraneous special characters while preserving cash tags and stock tickers as critical financial identifiers. This transformation converts raw tweets into cleaner, structured forms, establishing a consistent foundation for downstream processes like feature extraction, embedding generation, and model training
2) Bot and Spam Filtering: Dataset quality protection employs a dual strategy for bot and spam detection. A lightweight classifier trained on account metadata (including account age, follower-to-following ratios, and posting patterns) identifies potentially automated accounts. Simultaneously, rule-based heuristics detect repetitive spam patterns, such as accounts posting identical stock messages across multiple tickers within short timeframes.
3) Text Normalization: Contractions expand (e.g., “won’t” becomes “will not”) and text standardizes to lowercase. Emojis and emoticons convert to unique sentiment tokens like ‘smile’, preserving emotional context for learning algorithms.
4) Financial Jargon Normalization: A custom financial dictionary normalizes domain-specific slang and abbreviations. For instance, “BTFD” maps to “buy the dip”, and “ER” to “earnings report”. This ensures terminological consistency and helps models capture the intended meanings behind informal expressions.

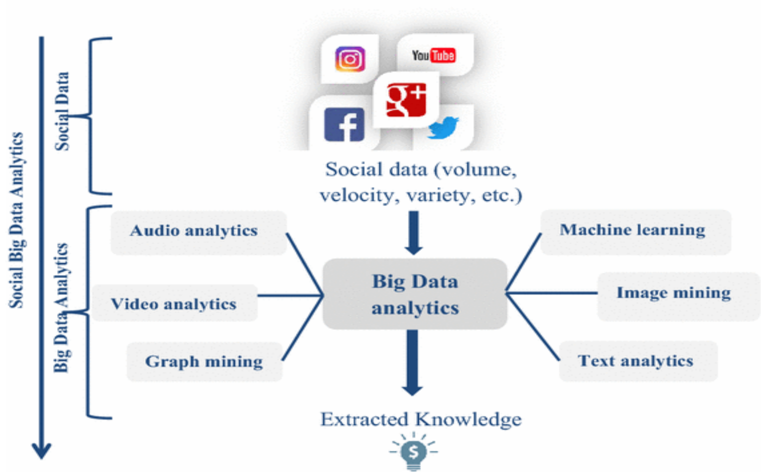

Fig. 2: Detailed Text Preprocessing Pipeline Addressing Specific Challenges in Financial Social Media Data.

3.2. Sentiment analysis ensemble

To mitigate limitations from single-model reliance and inherent inductive biases, we designed an ensemble framework integrating three distinct deep learning architectures. Each model produces probability distributions across three sentiment categories: Positive, Neutral, and Negative.

a) Bidirectional LSTM (Bi-LSTM): Bidirectional LSTMs have become standard in NLP tasks, including sentiment detection, part-of-speech tagging, and machine translation. Their effectiveness stems from capturing dependencies in both sequence directions, particularly valuable when word interpretation depends on both preceding and following context. Beyond text, Bi-LSTMs have proven successful in speech recognition, bioinformatics, and time-series forecasting.

The ensemble’s first component employs a Bidirectional LSTM. Tweets undergo tokenization and transformation into 300-dimensional GloVe embeddings serving as network input. These embeddings pass through a Bi-LSTM layer with 128 hidden units, enabling contextual information retention from both left-to-right and right-to-left sequences. Output feeds into a fully connected dense layer with softmax activation, yielding final sentiment probabilities.

b) Convolutional Neural Network (CNN): Following Kim’s approach [9], the CNN architecture identifies salient n-gram patterns in text. The model uses parallel one-dimensional convolutional layers with kernel sizes 3, 4, and 5, each generating 100 feature maps. Outputs undergo max-over-time pooling, extracting the most prominent signal from each map. Pooled features are concatenated and passed through a fully connected dense layer for final classification.

While CNNs demonstrate strong text categorization performance, they require substantial computational resources and large labelled datasets for effective training. Over-fitting presents a notable challenge, particularly with smaller datasets. Common mitigation techniques include dropout regularization, transfer learning, and data augmentation.

c) Fine-Tuned BERT: Fine-tuned BERT adapts the pre-trained Bidirectional Encoder Representations from Trans- formers model for specific tasks. BERT undergoes initial training on large text volumes using general objectives like masked word prediction and sentence order prediction. This training phase provides strong language structure and context understanding. The ensemble’s foundation uses a Transformer-based model, specifically ‘BERT-base-uncased‘. Financial domain adaptation involves a two-step fine-tuning:

1) Domain-Adaptive Pre-training: Base BERT undergoes continued pre-training on extensive, unlabeled financial news and tweet corpora via Masked Language Modeling. This helps the model internalize financial vocabulary and context.
2) Supervised Fine-tuning: The domain-adapted model undergoes fine-tuning on manually labeled sentiment datasets. A single linear classification layer sits atop the [CLS] token’s final embedding for this task.

d) Ensemble Fusion: Three base models’ probability vectors are concatenated, forming a 9-dimensional feature vector. This vector inputs to a Meta learner shallow Multi-Layer Perceptron with one hidden layer containing 16 neurons. This meta-learner trains to produce the final consolidated sentiment prediction. This stacking approach enables the system to learn optimal combinations of base classifier judgments.

3.3. Stock direction prediction module

The pipeline's final stage predicts next-day closing price direction (whether it rises or falls).

a) Feature Engineering: At each trading day's close, we construct a feature vector for each stock from two information sources:

- Sentiment Features: These features aggregate all stock-associated tweets from the preceding 24 hours. The feature set includes
- Average sentiment score, where positive = 1, neutral= 0, and negative = –1.
- Relative proportion of tweets in each sentiment category.
- Total tweet count, reflecting market attention and public engagement levels.
- Sentiment variability, measured using the standard deviation of sentiment scores.

These sentiment indicators provide structured representations of public opinion dynamics

- Technical Indicators: Calculated from historical OHLCV data. The feature set includes standard indicators:
- Moving Averages (SMA-10, EMA-20).
- Relative Strength Index (RSI-14).
- Moving Average Convergence Divergence (MACD).
- Bollinger Bands (upper and lower).
- 10-day realized volatility.

All features undergo z-score normalization before model input for consistency

b) Prediction Model Architecture: To effectively process these heterogeneous feature types, we employ a dual-branch neural network.

Table 1: Dataset Summary

Component	Size	Period	Source
Twitter	2.3M tweets	Jan 2020–Dec 2022	Twitter API v2.0
Stock prices	50 stocks	Jan 2020–Dec 2022	Yahoo Finance
News articles	180K	Jan 2020–Dec 2022	Reuters, Bloomberg
Labeled sentiment	100K	Various	Manual annotation

- Temporal Branch: This branch receives 20-day technical indicator sequences, processing them with an LSTM layer to learn from recent price trends and momentum.
- Static Branch: Current-day aggregated sentiment features pass through a standard dense layer.
- Fusion Layer: Temporal and static branch outputs concatenate. This unified representation passes through two fully connected layers using ReLU activation and dropout for regularization. The final output is a single neuron with sigmoid activation, providing the price increase probability.

The model trains on binary targets: 1 if 'Close (t+1) Close (t), 0 otherwise. To address the inherent imbalance between "up" and "down" market days, we use weighted binary cross-entropy loss.

4. Experimental Setup

4.1. Datasets

Table I summarizes the main data sources.

Labelled data combined expert annotation with quality-controlled crowd labels (measuring inter-annotator agreement and conducting spot checks). Weak labels from emojis, finance keywords, and news headlines expanded coverage.

4.2. Implementation

Key implementation details

- Python 3.8, PyTorch 1.11, Tensor Flow 2.8, and Hugging Face Transformers.
- Large-scale preprocessing using Apache Spark (PySpark); Parquet storage.
- Training on NVIDIA Tesla V100 GPUs; experiment tracking with Weights & Biases.

4.3. Evaluation protocol and tuning

We adopt time-respecting evaluation:

- Rolling-origin walk-forward validation (training on [t0, tk], validating on (tk, tk+1], testing on (tk+1, tk+2]) to simulate deployment.
- Hyperparameter tuning via Bayesian optimization (Op- tuna) over constrained ranges; early stopping on validation F1 for sentiment and direction accuracy for the predictor.
- Probability calibration (temperature scaling) applied on held-out slices to improve threshold stability across periods.

Representative ranges:

- LSTM hidden units: [64, 256], dropout: [0.2, 0.6]
- Learning rates: [1e-5, 1e-3]
- Batch sizes: [16, 64] for BERT; [64, 512] for lighter models

The final selected hyperparameters appear in Table II.

Table 2: Selected Hyper Parameter

Component	Parameter	Value
LSTM	Hidden units	128
LSTM	Dropout	0.5
CNN	Filter sizes	[3,4,5]
CNN	Feature maps	100
BERT	Learning rate	2e-5
BERT	Batch size	32
Training	Epochs	50

4.4. Metrics

1) Classification Metrics: Reporting includes Accuracy, Precision, Recall, and F1-score
- Accuracy measures correctly predicted instances as a fraction of total samples. While providing an overall performance perspective, it can mislead in imbalanced datasets, as the majority-class bias can still yield high accuracy.
- Precision assesses positive prediction reliability, defined as true positives among all positive predictions. High precision indicates rare false positive occurrences—crucial in domains like fraud detection or spam filtering.
- Recall evaluates completeness by measuring the proportion of actual positives correctly identified. High recall proves essential in applications like medical screening, where missing true positives carries severe consequences.
- F1-score balances precision and recall through its harmonic mean. This metric proves particularly informative with class imbalance, where accuracy alone fails to reflect the false positive-false negative trade-off.
2) Financial Simulation Metrics: To evaluate prediction effectiveness in real trading contexts, we track cumulative return, annualized return, Sharpe ratio, and maximum drawdown. These metrics collectively provide insights into profitability, risk-adjusted performance, and downside exposure.
3) Operational Metrics: From system efficiency perspectives, prediction latency and throughput monitoring ensure the model delivers timely and scalable outputs in production environments.

5. Results

5.1. Sentiment classification

Table 3 presents held-out performance. BERT proves the strongest single classifier; the ensemble improves both mean accuracy and stability.

Table 3: Sentiment classification results

Model	Accuracy	Precision	Recall	F1
LSTM only	0.823	0.821	0.819	0.820
CNN only	0.798	0.795	0.792	0.793
BERT only	0.856	0.854	0.851	0.852
Proposed ensemble	0.873	0.871	0.869	0.870
Baseline (SVM)	0.721	0.718	0.715	0.716

Table 4: Stock Prediction and Backtest Metrics

Metric	Value
Direction accuracy	82.1%
Precision (up)	0.834
Precision (down)	0.808
Average return	12.3%
Sharpe ratio	1.67
Max drawdown	8.2%

5.2. Stock prediction and backtest

The integrated model achieves 82.1% direction accuracy across tested tickers. We simulate a daily strategy taking long exposure when the model predicts ”up” and holding cash otherwise.

5.3. Ablation and robustness

We evaluate component contributions and robustness:
- Removing sentiment features decreases direction accuracy by approximately 7.4 points, confirming their incremental value beyond technical.
- Using only BERT sentiment (without ensemble) lowers mean accuracy by 1.7 points and increases variance during volatile periods.
- Excluding technical indicators reduces average return by roughly 4.8%, indicating complementarity between price-derived and social signals.
- Probability calibration improves threshold stability, reducing false positives following regime shifts (e.g., earnings seasons).

5.4. Error analysis and calibration

Qualitative inspection reveals common failure modes: sarcasm, ambiguous ticker mentions (e.g., common words overlapping with tickers), and event clustering (multiple firms reporting earnings simultaneously). Calibrated probabilities reduce overconfident errors, improving decision thresholds for trading rules.

6. Operational Consideration

6.1. Latency and throughput

In our configuration:
- Ingestion and preprocessing for 100-ticker batches complete in less than 2 seconds per minute of tweets on a 4-GPU node.

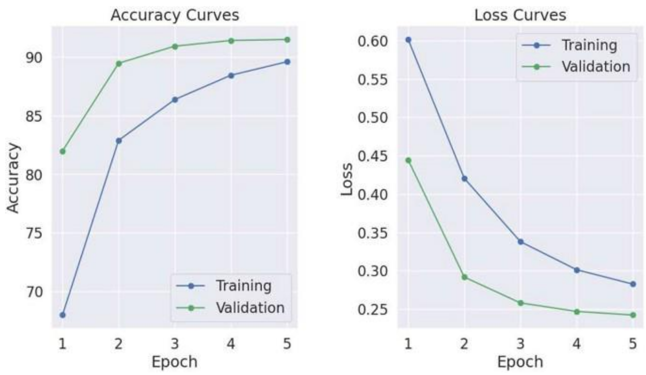

Fig. 3: Training and Validation Curves for Sentiment Models.

Fig. 4: Cumulative Returns Comparing Simulated Strategy Versus Buy-and-Hold.

- BERT inference dominates computation; quantization, knowledge distillation, or lightweight domain-adapted encoders can reduce production latency.

6.2. MLOps and drift handling

Deployment includes rolling retraining, drift detection on input distributions, and shadow evaluation for new models. Canary releases with capital limits mitigate deployment risk. Feature stores ensure consistent definitions between training and inference.

6.3. Computational cost

Fine-tuning BERT components required approximately 12 GPU-hours per run on V100. Full pipeline experiments (multiple seeds and hyperparameter searches) consumed several hundred GPU-hours.

7. Discussion

Experimental findings demonstrate that combining social media sentiment with conventional technical indicators noticeably improves short-term stock market prediction accuracy. This suggests online market sentiment carries signals complementing historical price-based features.

a) Interpretation of Results: The sentiment ensemble's enhanced performance stems from the underlying models' diverse strengths. Bi-LSTM captures sequential opinion flows, detecting gradual tone shifts across post timelines. CNN proves more sensitive to high-impact expressions like "sell-off" or "unexpected gains", while BERT contributes by interpreting complex linguistic structures and contextual nuances. The meta-learner integrates these perspectives, yielding more reliable outcomes than any single model.

The observed 7.4 percent predictive accuracy improvement after adding sentiment to the technical baseline proves particularly significant. It demonstrates that social media reactions often precede or amplify price movements, offering additional insights not contained in market data alone. This aligns with the intuition that investors' immediate responses to news, earnings, or rumors surface on platforms like Twitter before full absorption into asset prices.

Beyond empirical performance, the findings also align with principles of market efficiency and behavioural finance, suggesting that investor sentiment reflected on Twitter contributes to short-term inefficiencies in asset pricing. This supports the notion that emotional and cognitive biases, observable through social data, influence market behaviour before full price adjustment. The framework thus not only aids forecasting but also offers a computational lens for understanding investor psychology and crowd behaviour in digital financial ecosystems.

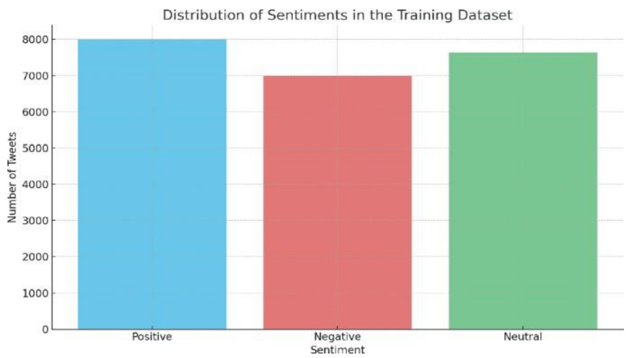

Fig. 5: Ablation Study: Directional Accuracy When Excluding Feature Groups.

8. Ethics, Privacy, and Bias

We utilize only publicly available tweets under platform terms. User identifiers undergo hashing when unnecessary for analysis. Bot detection removes likely automated accounts to reduce manipulation. We avoid collecting or inferring protected attributes. Nonetheless, social data can encode biases and face coordinated campaigns; monitoring, audit trails, and human oversight prove essential before using signals in live trading.

9. Reproducibility and Data Availability

We provide hyperparameter ranges, seeds, and detailed experiment logs. Requirements files and container specifications are available on request. Raw Twitter data remains subject to platform terms and cannot be redistributed; we supply scripts to re-collect tweets given access tokens and filters.

10. Limitations

Key limitations include: (i) sample bias in social media users; (ii) performance sensitivity to structural market changes; (iii) simplified backtest assumptions (transaction costs, slippage, borrow constraints), and (iv) susceptibility to adversarial narratives and event clustering. Future work should stress-test under extreme regimes and expand risk-aware decision rules.

Additionally, the backtesting framework assumes frictionless trading, omitting transaction costs, slippage, and liquidity constraints, which could affect real-world profitability. Incorporating such factors in future simulations will provide a more realistic assessment of financial feasibility. Further, order book depth and execution latency may introduce additional variance in practical deployment.

11. Conclusion and Future Work

We presented an integrated pipeline using Twitter sentiment, processed by deep model ensembles, combined with technical indicators to forecast short-term stock direction. Results indicate consistent gains over baselines and underscore the carefully engineered social signals' value.

Although the ensemble approach enhances accuracy, interpretability remains a challenge. Emerging tools such as attention visualization, SHAP, and LIME offer pathways to better understand model decision-making. Future iterations could integrate such techniques to highlight influential tweet features or sentiment drivers, enhancing transparency and trust in model outputs.

Future work includes:

1) Expanding to multi-modal inputs (news, filings) and quantifying marginal utility.
2) Graph-based modelling of ticker and user interactions.
3) Distilled or sparse transformers for low-latency streaming inference.
4) Explain ability modules surfacing influential tweets and features per prediction, improving trust and debug ability.

References

[1] Bollen, H. Mao, and X. Zeng, "Linking public mood on Twitter with movements in the stock market," *J. Comput. Sci.*, vol. 2, no. 1, pp. 1–8, 2011. https://doi.org/10.1016/j.jocs.2010.12.007.

[2] X. Zhang, H. Fuehres, and P. A. Gloor, "Forecasting market signals using Twitter discussions," *Procedia–Social Behav. Sci.*, vol. 26, pp. 55–62, 2018. https://doi.org/10.1016/j.sbspro.2011.10.562.

[3] A. Kumar and V. Ravi, "Applications of text mining in finance: A survey," *Knowl.-Based Syst.*, vol. 114, pp. 128–147, 2016. https://doi.org/10.1016/j.knosys.2016.10.003.

[4] D. Shah, H. Isah, and F. Zulkernine, "A taxonomy and review of stock prediction methodologies," *Int. J. Financ. Stud.*, vol. 7, no. 2, p. 26, 2019. https://doi.org/10.3390/ijfs7020026.

[5] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Deep bidirectional transformers for language understanding," in *Proc. NAACL- HLT*, 2019, pp. 4171–4186. https://doi.org/10.18653/v1/N19-1423.

[6] Y. Li, S. Pan, K. Zhang, Y. Yin, and Q. Chen, "Financial sentiment learn- ing for stock return prediction," *IEEE Access*, vol. 8, pp. 78899–78908, 2020.

[7] T. Chen, R. Xu, Y. He, and X. Wang, "Enhancing sentiment analysis via sentence-type features with BiLSTM-CRF and CNN," *Expert Syst. Appl.*, vol. 72, pp. 221–230, 2021. https://doi.org/10.1016/j.eswa.2016.10.065.

[8] S. Hochreiter and J. Schmidhuber, "Long short-term memory networks, "*Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997. https://doi.org/10.1162/neco.1997.9.8.1735.

[9] Y. Kim, "CNNs for sentence-level text classification," in *Proc. EMNLP*, 2014, pp. 1746–1751. https://doi.org/10.3115/v1/D14-1181.

[10] A. Vaswani *et al.*, "Attention mechanisms in neural networks: The Transformer model," in *Proc. NeurIPS*, 2017, pp. 5998–6008.

[11] R. Socher *et al.*, "Recursive deep learning for sentiment treebank analysis," in *Proc. EMNLP*, 2013, pp. 1631–1642. https://doi.org/10.18653/v1/D13-1170.

[12] M. E. Peters *et al.*, "ELMo: Deep contextualized word representations," in *Proc. NAACL-HLT*, 2018, pp. 2227–2237. https://doi.org/10.18653/v1/N18-1202.

[13] Z. Yang *et al.*, "Hierarchical attention networks for document-level text classification," in *Proc. NAACL-HLT*, 2016, pp. 1480–1489.

[14] X. Li, H. Xie, L. Chen, J. Wang, and X. Deng, "Impact of news sentiment on stock returns," *Knowl.-Based Syst.*, vol. 69, pp. 14–23, 2014. https://doi.org/10.1016/j.knosys.2014.04.022.

[15] P. C. Tetlock, "Investor sentiment and the influence of media on stock markets," *J. Finance*, vol. 62, no. 3, pp. 1139–1168, 2007. https://doi.org/10.1111/j.1540-6261.2007.01232.x.

[16] T. O. Sprenger, A. Tumasjan, P. G. Sandner, and I. M. Welpe, "Infor- mation value of Twitter stock microblogs," *Eur. Financ. Manag.*, vol. 20, no. 5, pp. 926–957, 2014. https://doi.org/10.1111/j.1468-036X.2013.12007.x.

[17] H. Yu, J. Nartea, C. Gan, and L. J. Yao, "Profitability of technical trading rules in Southeast Asian markets," *Int. Rev. Econ. Finance*, vol. 25, pp. 356–371, 2013. https://doi.org/10.1016/j.iref.2012.07.016.

[18] B. Weng, F. F. Ahmed, and F. M. Megahed, "Short-term stock movement prediction using heterogeneous data sources," *Expert Syst. Appl.*, vol. 79,pp. 153–163, 2017. https://doi.org/10.1016/j.eswa.2017.02.041.

[19] L. Kristoufek, “Drivers of Bitcoin price identified through wavelet coherence,” *PLoS ONE*, vol. 10, no. 4, e0123923, 2015. https://doi.org/10.1371/journal.pone.0123923.
[20] M. Ballings, D. Van den Poel, N. Hespeels, and R. Gryp, “Comparative evaluation of classifiers for predicting stock price direction,” *Expert Syst. Appl.*, vol. 42, no. 20, pp. 7046–7056, 2015. https://doi.org/10.1016/j.eswa.2015.05.013.
[21] Z. Yang et al., “FinBERT-X: Domain-Specific Transformer Models for Financial Sentiment Analysis,” *Expert Systems with Applications*, vol. 213, 2023.
[22] L. Chen and J. Wang, “RoBERTa-Finance: Improved Contextual Representations for Financial NLP,” *IEEE Access*, vol. 11, pp. 55924–55937, 2023.