



Market Mind Analyzer: An AI-Driven Framework for Financial Sentiment, Market Trend Analysis, and Predictive Decision Support

Dikshita¹, Vijay Gautam², Dr B K Verma³

Department of CSE-AI&DS,

Panipat Institute of Engineering & Technology, Samalkha, India^{1,2,3}

Mail ID: hod.cse.ai-ds@piet.co.in

ABSTRACT--Financial markets are influenced by historical price behavior, technical indicators, macroeconomic conditions, corporate events, financial news, and investor sentiment. The high volume and velocity of financial information make manual analysis increasingly difficult. This paper proposes Market Mind Analyzer (MMA), an artificial intelligence-driven decision-support framework that integrates financial market data with natural language processing (NLP), sentiment analysis, technical indicators, machine learning, and explainable visualization. The proposed framework collects financial news and market information, preprocesses textual and numerical data, extracts sentiment using a domain-adapted language model such as FinBERT, calculates technical indicators, and combines these heterogeneous features for market trend classification. The system produces a composite Market Mind Score representing the prevailing positive, neutral, or negative market state and generates an interpretable dashboard containing sentiment trends, technical signals, volatility information, and predicted market direction. Unlike systems based exclusively on historical prices or textual sentiment, MMA combines behavioral and quantitative signals in a unified architecture. The paper presents the system architecture, feature-engineering methodology, prediction strategy, evaluation metrics, and limitations. Numerical experimental results are intentionally not fabricated; the proposed evaluation protocol is provided for execution on a selected dataset.

Keywords: Financial Market Prediction, Sentiment Analysis, FinBERT, Natural Language Processing, Machine Learning, Technical Indicators, Market Intelligence, Explainable AI, Stock Market, Decision Support System.

INTRODUCTION

Financial markets represent complex adaptive systems in which prices continuously respond to economic information, company-specific events, investor expectations, market liquidity, and collective behavior. Traditional financial analysis generally combines fundamental analysis, technical analysis, and market knowledge. However, the increasing volume of digital financial information has created a requirement for automated analytical systems capable of processing both structured and unstructured data.

The growth of online financial news, social media, corporate announcements, analyst reports, and investor discussions has created a large textual information environment. Such information may influence investor expectations and consequently affect market activity. Tetlock demonstrated a relationship between media pessimism and subsequent market behavior [1]. Loughran and McDonald showed that



general-purpose sentiment dictionaries can misclassify financial terminology and proposed domain-specific financial lexicons [2].

Transformer-based language models have substantially improved contextual text understanding. BERT introduced bidirectional contextual representations for downstream NLP tasks [3], while FinBERT adapted transformer-based modeling to financial sentiment analysis [4]. Machine-learning and deep-learning methods including support vector machines, random forests, gradient boosting, neural networks, and LSTM models have also been widely investigated for financial forecasting [5], [6].

This paper proposes Market Mind Analyzer (MMA), an integrated AI-based framework combining financial-news sentiment, historical market prices, technical indicators, volatility analysis, machine-learning-based trend classification, a composite Market Mind Score, and explainable visualization.

The main contributions are: (1) a unified architecture for textual and numerical financial information; (2) a domain-specific sentiment component; (3) a composite Market Mind Score; (4) a machine-learning trend-prediction module; and (5) an explainable market-intelligence dashboard.

LITERATURE REVIEW

A. Financial Sentiment and Market Behavior

Tetlock analyzed financial media content and reported that pessimistic media content was associated with market pressure and trading activity [1]. Antweiler and Frank investigated Internet stock message boards and found that online discussions contain information relevant to stock-market behavior [7]. These studies established an important foundation for using textual information as an additional market signal.

B. Financial Text Processing

Loughran and McDonald demonstrated that conventional sentiment dictionaries are not sufficiently reliable for financial documents because financial terminology has domain-specific meanings [2]. FinBERT subsequently provided a transformer-based approach specialized for financial sentiment classification [4].

C. Machine Learning for Financial Prediction

Financial prediction research has investigated SVM, random forests, neural networks, LSTM, and hybrid models. Recent systematic reviews identify sentiment, historical returns, momentum, volatility, and technical indicators as important candidate features for financial prediction [5], [6].

D. Hybrid and Explainable Approaches

Recent research increasingly combines textual sentiment with numerical market variables and deep learning. Decision-fusion approaches have also been explored for combining predictions from multiple models [8]. Recent studies further emphasize the importance of evaluating financial usefulness rather than relying only on predictive accuracy [9].

Approach	Primary Data	Technique	Strength	Limitation
Technical analysis	Price/volume	Indicators	Simple, interpretable	Ignores text
Dictionary sentiment	Financial text	Lexicon	Efficient	Limited context
ML forecasting	Market data	SVM/RF/ANN	Nonlinear modeling	Non-stationarity
LSTM	Time series	LSTM	Temporal	Data/computation



			dependencies	
FinBERT	Financial text	Transformer	Context-aware	Computational cost
FinBERT + LSTM	News + price	Hybrid DL	Multimodal	Complexity
Market Mind Analyzer	News + price + indicators	Fusion + ML	Integrated, explainable	Requires aligned data

RESEARCH GAP

Existing approaches frequently treat sentiment analysis and numerical market prediction as separate problems. Other systems emphasize prediction accuracy without sufficient attention to interpretability or financial utility. A unified framework is therefore needed to combine textual sentiment, market behavior, technical signals, prediction, and explainability. MMA addresses this gap through a multimodal decision-support architecture.

DATA ACQUISITION

The system can use historical OHLCV data, financial news, company information, market indices, technical indicators, and sentiment labels. For experimental work, publicly accessible historical market data can be aligned with a labeled financial-news dataset. All records must be timestamp-aligned to prevent future-information leakage.

Data Category	Example Variables	Purpose
OHLCV	Open, High, Low, Close, Volume	Price analysis
Financial news	Headline, article, timestamp	Sentiment
Company information	Company, sector, ticker	Entity mapping
Market indices	NIFTY, SENSEX, S&P 500	Market context
Technical indicators	RSI, MACD, SMA, EMA	Trend identification
Volatility	Rolling SD, ATR	Risk analysis
Sentiment	Positive/neutral/negative probabilities	Behavioral signal

DATA PREPROCESSING

Financial text can contain ticker symbols, abbreviations, numerical values, company names, URLs, duplicate articles, and domain-specific terminology. The preprocessing pipeline includes duplicate removal, URL and HTML removal, text normalization, timestamp normalization, entity/ticker extraction, missing-value treatment, outlier analysis, and temporal alignment. Extreme market observations should be investigated rather than automatically deleted because financial markets naturally contain extreme events.

SENTIMENT ANALYSIS MODULE

For a news item x , a financial sentiment model produces $P(y|x) = [P_{\text{positive}}, P_{\text{neutral}}, P_{\text{negative}}]$. The sentiment score is defined as $S_t = P_{\text{positive}} - P_{\text{negative}}$. A positive value indicates positive textual sentiment, a negative value indicates negative sentiment, and a value close to zero indicates neutrality.



A weighted daily sentiment score is calculated as $S_d = \sum(w_i S_i) / \sum w_i$, where w_i can represent article relevance, recency, source reliability, and model confidence.

TECHNICAL FEATURE ENGINEERING

Simple Moving Average (SMA) is defined as $SMA_t = (1/n) \sum P_{(t-i)}$. Relative Strength Index is $RSI = 100 - 100/(1+RS)$. MACD is $MACD = EMA_{12} - EMA_{26}$. Daily return is $R_t = (P_t - P_{(t-1)})/P_{(t-1)}$. Rolling volatility is estimated using the standard deviation of returns. These features represent trend, momentum, and risk.

FEATURE FUSION

The fused feature vector is represented as $X_t = [S_t, R_t, RSI_t, MACD_t, SMA_t, V_t, \sigma_t, \dots]$. The fused representation is supplied to machine-learning models. This allows the system to examine whether textual sentiment adds predictive information beyond conventional market variables.

MARKET DIRECTION PREDICTION

MMA formulates the initial task as three-class classification: $Y_{(t+1)} \in \{\text{Bullish, Neutral, Bearish}\}$. A target can be defined using a return threshold θ : Bullish when $R_{(t+1)} > \theta$, Neutral when $-\theta \leq R_{(t+1)} \leq \theta$, and Bearish when $R_{(t+1)} < -\theta$. Candidate models include Logistic Regression, Random Forest, XGBoost, and LSTM. A transformer-based temporal model may be investigated in future work.

MARKET MIND SCORE

The Market Mind Score (MMS) combines sentiment, returns, momentum, and volatility-adjusted signals: $MMS_t = w_s S_t + w_r R_t + w_m M_t + w_v V_t$, subject to $w_s + w_r + w_m + w_v = 1$. The weights and decision thresholds should be learned or validated on historical data rather than selected only for presentation.

MMS Range	Market State	Interpretation
0.60 to 1.00	Strong Bullish	Strong positive combined signal
0.20 to 0.59	Bullish	Positive market condition
-0.19 to 0.19	Neutral	Mixed/uncertain condition
-0.59 to -0.20	Bearish	Negative market condition
-1.00 to -0.60	Strong Bearish	Strong negative combined signal

EXPERIMENTAL DESIGN

A chronological dataset should be divided into training, validation, and testing subsets. A possible configuration is 70% training, 15% validation, and 15% testing. Unlike random shuffling, chronological splitting is appropriate for time-dependent prediction because it reduces temporal leakage.

Parameter	Proposed Setting
Problem	Market direction classification
Classes	Bullish, Neutral, Bearish
Text model	FinBERT
Baseline 1	Logistic Regression
Baseline 2	Random Forest



Baseline 3	XGBoost
Deep model	LSTM
Validation	Walk-forward validation
Primary metric	F1-score
Secondary metrics	Accuracy, Precision, Recall, ROC-AUC
Financial metrics	Sharpe ratio, cumulative return, maximum drawdown
Leakage control	Chronological splitting

RESULTS AND DISCUSSION

The experimental results must be generated from the selected dataset using identical temporal splits for all models. The following table defines the comparison rather than reporting fabricated performance values.

Model	Price	Technical	Sentiment	Purpose
Logistic Regression	✓	✓	✗	Baseline
Random Forest	✓	✓	✗	Nonlinear baseline
XGBoost	✓	✓	✓	Feature-fusion model
LSTM	✓	✓	✗	Temporal baseline
FinBERT + ML	✓	✓	✓	Text-enhanced model
MMA	✓	✓	✓	Integrated framework

Existing studies support the design rationale. FinBERT has demonstrated strong financial sentiment-analysis capability [4]. Research combining financial sentiment with sequential market models indicates that textual sentiment can provide useful information for stock-movement prediction [10]. Recent IEEE literature also identifies hybrid sentiment and market-data approaches as an active research direction [5], [6]. These findings motivate empirical testing of MMA, but they do not constitute experimental results for the proposed implementation.

FUTURE WORK

Future work can incorporate large language models for richer financial-document reasoning, multilingual financial NLP for Indian-language content, sector-specific models, portfolio-level Market Mind Scores, graph neural networks for inter-stock relationships, real-time streaming architectures, and user-centered evaluation of explainable signals. Graph-transformer approaches that integrate BERT-derived sentiment represent another promising research direction.

CONCLUSION

This paper presented Market Mind Analyzer, an AI-driven framework integrating financial sentiment, historical market data, technical indicators, machine learning, and explainable visualization. The framework moves beyond isolated price prediction toward multimodal market intelligence. Existing research supports the relevance of media sentiment, financial-domain language modeling, and hybrid



machine-learning approaches. The proposed system should be validated through chronological testing, ablation studies, statistical significance testing, transaction-cost-aware backtesting, and evaluation across multiple market regimes. MMA is therefore positioned as an academic decision-support framework rather than an autonomous investment system.

REFERENCES

- [1] P. C. Tetlock, "Giving content to investor sentiment: The role of media in the stock market," *The Journal of Finance*, vol. 62, no. 3, pp. 1139–1168, 2007, doi: 10.1111/j.1540-6261.2007.01232.x.
- [2] T. Loughran and B. McDonald, "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," *The Journal of Finance*, vol. 66, no. 1, pp. 35–65, 2011, doi: 10.1111/j.1540-6261.2010.01625.x.
- [3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186, doi: 10.18653/v1/N19-1423.
- [4] D. Araci, "FinBERT: Financial sentiment analysis with pre-trained language models," *arXiv:1908.10063*, 2019.
- [5] "A systematic literature review for stock price prediction with machine learning," in *Proc. 11th Int. Conf. Communication and Signal Processing (ICCSP)*, 2025, doi: 10.1109/ICCSP64183.2025.11088479.
- [6] "A comprehensive review of machine learning for financial market prediction methods," in *Proc. Int. Conf. Emerging Smart Computing and Informatics (ESCI)*, 2023, doi: 10.1109/ESCI56872.2023.10099791.
- [7] W. Antweiler and M. Z. Frank, "Is all that talk just noise? The information content of Internet stock message boards," *The Journal of Finance*, vol. 59, no. 3, pp. 1259–1293, 2004.
- [8] "Decision fusion for stock market prediction: A systematic review," *IEEE Access*, vol. 10, pp. 81364–81379, 2022, doi: 10.1109/ACCESS.2022.3195942.
- [9] "A systematic review of stock market forecasting utilizing machine learning and deep learning," in *Proc. IEEE 16th Control and System Graduate Research Colloquium (ICSGRC)*, 2025, doi: 10.1109/ICSGRC65918.2025.11159810.
- [10] T. Jiang and A. Zeng, "Financial sentiment analysis using FinBERT with application in predicting stock movement," *arXiv:2306.02136*, 2023.
- [11] M. Al Ridhawi, M. H. Ali, and H. Al Osman, "Stock market prediction using Node Transformer architecture integrated with BERT sentiment analysis," *arXiv:2603.05917*, 2026.
- [12] K. Kirtac and G. Germano, "Enhanced financial sentiment analysis and trading strategy development using large language models," in *Proc. Workshop on Financial Technology and Natural Language Processing*, 2024, pp. 1–10, doi: 10.18653/v1/2024.wassa-1.1.