Iranian Journal of Finance

Print ISSN 2676-6337
Online ISSN 2676-6345

Forecasting Market Closing Direction: A Novel Data Fusion of Deep News Embeddings and Technical Indicators with Gradient Boosting

Fatemeh Zare Baghiabad *

*Corresponding author, Assistant prof., Department of Industrial Engineering, Faculty of Engineering, Ardakan University, Ardakan, Iran. (Email: f.zare@ardakan.ac.ir)

Marzieh Karimi

Assistant prof., Department of Industrial Engineering, Faculty of Engineering, Ardakan University, Ardakan, Iran. (Email: m.karimi@ardakan.ac.ir)

Iranian Journal of Finance, 2026, Vol. 10, No.2, pp. 62-84.
Publisher: Iran Finance Association
doi: https://doi.org/10.30699/ijf.2026.570396.1570
Article Type: Original Article

Received: January 07, 2026
Received in revised form: February 27, 2026
Accepted: March 22, 2026
Published online: April 01, 2026

Abstract

Accurately predicting stock market movements remains a complex engineering challenge due to inherent non-linearity and volatility. This study addresses this problem by developing a novel artificial intelligence framework that fuses deep learning-based natural language processing with technical analysis for financial forecasting. Its primary engineering application is predicting the daily directional movement of the Tehran Stock Exchange total index, a critical task for automated trading systems and risk management. The core AI contribution is a hybrid architecture that generates contextual text embeddings from Persian news using a Bidirectional Encoder Representations from Transformers

(BERT) model and integrates them with technical indicators. These features are processed by a soft-voting ensemble of powerful gradient-boosting algorithms, namely eXtreme Gradient Boosting (XG-Boost) and Light Gradient Boosting Machine (LightGBM). Rigorously validated through an extensive walk-forward testing procedure across 331 temporal windows, the model achieved impressive performance metrics: 84.0% accuracy, 86.1% F1 Score, 87.0% precision, and 85.2% recall. These notable results demonstrate the significant value of domain-specific natural language processing in non-English financial contexts. Furthermore, this work establishes gradient boosting ensembles as a highly efficient, high-performance alternative to Long Short-Term Memory (LSTM) networks for financial forecasting. The proposed methodology demonstrates robust predictive capability for directional movement forecasting in emerging markets.

Keywords: Stock Market Forecasting, Persian Natural Language Processing, Bidirectional Encoder Representations from Transformers Embeddings, Gradient Boosting Ensemble, Multimodal Data Fusion

JEL Classification: G21, C53, C45, C38

Introduction

Forecasting fluctuations in the financial market is one of the most difficult yet interesting applications of modern computational intelligence (Oluwatobiloba, 2025). The noise, non-stationarity, and sophisticated volatility that characterize historical market data have resulted in an unachievable precision scale in forecasting. However, the introduction of deep learning has spurred the formation of a new paradigm in computational intelligence, as it provides robust and flexible features to characterize intricate, nonlinear signals generated from financial time series. As summarized in (Ohliati & Yuniarty, 2024), architectures such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM), new neural network architectures–such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks–have shown great promise in capturing temporal dependence, removing the requirements and needs of earlier statistical and machine learning models. This work establishes that models like LSTMs are particularly well-suited for sequential data, a finding corroborated by studies such as Li (2024), which systematically applied these deep neural networks to price prediction tasks.

A significant evolution in this field has been the move beyond purely numerical time-series data. Recognizing that market prices are not only a function of past prices but also of new information, researchers began integrating unstructured textual data from news articles and social media. The critical role of Natural Language Processing (NLP) in this endeavor is thoroughly explored in Takale (2024), which highlights methodologies for extracting sentiment and semantic meaning from financial text. The pinnacle of this approach is exemplified by the work of Hassan et al. (2022). This study directly aligns with the core methodology of the present research, as it demonstrates that an ensemble model based on LSTM networks, when fed high-quality, domain-specific word embeddings from Fin-BERT, can significantly enhance predictive reliability by effectively fusing quantitative data with qualitative news sentiment.

The landscape of deep learning applications in finance is diverse, with studies exploring a range of architectures and domains. For instance,(Ti, 2024) provides valuable empirical comparisons of different models, helping to delineate their relative strengths and weaknesses. Beyond direct price prediction, the research scope has expanded to include a broader analysis of market microstructure, as seen in Tan (2024), which uses these algorithms to decode complex patterns in trader behavior. Furthermore, the entire pipeline from data ingestion to model deployment is critical, a point emphasized by Zhang et al. (2024), which underscores the importance of robust data preprocessing and feature engineering as a prerequisite for successful deep learning models.

Recent advancements continue to underscore the pivotal role of news sentiment and alternative data in forecasting. The study by Mao et al. (2025) introduces an explainable deep learning framework specifically designed to forecast crude oil futures price volatility by integrating news sentiment, highlighting the value of model interpretability alongside predictive power in complex commodity markets. Similarly, the work of Gurgul et al. (2025)

demonstrates the efficacy of deep learning and NLP in cryptocurrency forecasting by fusing financial, blockchain, and social media data, highlighting the necessity of multimodal data fusion in highly volatile, sentiment-driven asset classes.

Despite these advancements, a clear research gap remains, which this study aims to address. While the LSTM-based ensemble with Fin-BERT represents a significant step forward, its application is typically confined to markets with extensive English-language news coverage and pre-trained models like Fin-BERT. The application of a similarly sophisticated NLP-driven ensemble model to other linguistic and market contexts, such as the Iranian stock market with its Persian-language news flow, remains underexplored. Furthermore, combining deep news embeddings with a wide array of technical indicators in a powerful gradient-boosting ensemble, either in place of or alongside LSTMs, presents a compelling alternative architecture.

Therefore, building upon the foundational work reviewed in Ohliati & Yuniarty (2024) and inspired by the news-integration framework of Hassan et al. (2022), this study proposes a novel hybrid forecasting framework. We integrate advanced feature engineering from technical analysis, as discussed in the comparative literature, with deep contextual embeddings of Persian-language news articles. By employing a robust ensemble of gradient-boosting models (XGBoost and LightGBM) validated through a rigorous walk-forward analysis, this research seeks to develop a reliable model for predicting the directional movement of the Iranian stock market, thereby contributing to the broader field of financial big data analysis, as outlined in the provided literature.

Research Methodology

This study introduces a novel hybrid forecasting framework that integrates deep learning-based news embeddings with traditional financial technical analysis to predict the directional movement of the Iranian stock market. The methodology is designed to capture the complex interplay between quantitative market data and qualitative news sentiment. The entire pipeline consists of four main stages: 1) Data Collection and Fusion, 2) Advanced Feature Engineering, 3) Ensemble Model Development, and 4) Rigorous Time-Series Validation. The process begins with raw financial data and news articles, which undergo sophisticated preprocessing and feature engineering. Technical indicators are calculated from price data, while news texts are converted into dense vector representations using a Persian BERT model. Features are scaled and reduced via PCA before being fed into an ensemble model (XG-Boost and Light-GBM) trained and evaluated using a rigorous walk-forward validation scheme to predict daily market direction.

This study leverages several key computational techniques. The Bidirectional Encoder Representations from Transformers (BERT) model, a state-of-the-art transformer-based architecture (Devlin et al., 2019), was employed to generate contextual embeddings from Persian news text due to its superior ability to capture semantic nuances in language. For predictive modelling, we utilized powerful gradient-boosting frameworks: XG-Boost (Chen & Guestrin, 2016) and LightGBM (Ke et al., 2017), which were selected for their proven high efficiency, accuracy in handling structured tabular data, and robustness against overfitting. Finally, Principal Component Analysis (PCA) (Sarkar, 2023) was applied for dimensionality reduction to mitigate multicollinearity among the numerous technical indicators and enhance model generalization.

Data Sources and Temporal Scope

The analysis covers a significant five-year period from December 22, 2018, to September 21, 2025 (corresponding to 1 Dey 1397 to 31 Shahrivar 1404 in the Persian calendar). This period includes diverse market regimes, providing a robust testbed for the model. Two primary data sources were integrated:

- Financial Data: The daily closing price of the Tehran Stock Exchange's total index was obtained directly from the exchange's official sources. The open price was derived by shifting the previous day's closing price.

- News Data: A comprehensive corpus of 3,210 news articles was collected via targeted web scraping from https://www.bourse24.ir/, a prominent Persian-language financial news platform. The corpus includes relevant political, economic, domestic, and international news that could influence investor sentiment in the Iranian market.

To ensure the absence of look-ahead bias, a strict timing was applied to the collection and matching of news data. News was filtered by date and local time in Iran. The timing rule was as follows: to predict trading day t, only news published from after the market closes on day t-1 (12:30) to before the market reopens on day t (9:00) was collected. This ensures that the model uses only information available at the time of prediction (before trading began). Then, according to equation (4), all news articles belonging to that particular day, after being converted into embedding vectors, were averaged into a single daily news embedding vector to be used as the input signal for that day.

Constructing the News Sentiment Embedding

To transform unstructured textual data into a quantitative signal, we employed a state-of-the-art Natural Language Processing (NLP) technique:

- Text Embedding with a Persian BERT Model: Each news article was processed using a pre-trained `BERT` model specifically optimized for the Persian language (`HooshvareLab/bert-fa-base-uncased`). This transformer-based model generates context-aware representations of text.
- From Words to a Daily Vector: For each article, the text was tokenized and passed through the BERT model. We extracted token embeddings and computed a fixed-length 768-dimensional vector for the entire article by applying mean pooling to the hidden states. This vector encapsulates the semantic meaning of the news.
- Aggregation to a Daily Signal: For each trading day, all article vectors were averaged to create a single Daily News Embedding Vector. This

vector serves as a dense, holistic representation of the aggregate news sentiment and information load for that day.

Target Variable Definition and Data Fusion

The predictive target is the market index's daily directional movement at close. The target variable y_t was formulated as a binary classification problem: $y_t = 1 \text{ if } \mathrm{Close_t} > \mathrm{Open_t}$ (Uptrend) and $y_t = 0 \text{ if } \mathrm{Close_t} \leq \mathrm{Open_t}$ (Downtrend). The final dataset was constructed by temporally aligning the financial data (open/close prices) with the corresponding daily news embedding vector, resulting in a unified time-series dataset.

Technical Analysis and Features

A suite of technical indicators was calculated from the historical price series using the TA-Lib library, capturing trends, momentum, and volatility:

- Trend: Simple Moving Averages (SMA: 3,5,10,20,50), Exponential Moving Averages (EMA: 3,5,10,20,50), MACD.
- Momentum: Relative Strength Index (RSI: 6,14,21), Rate of Change (ROC: 5,10), Momentum (5,10).
- Volatility: Bollinger Bands (including bandwidth position), and standard deviation (10,20-day periods)

News-Derived Features

Beyond using the raw embedding, we extracted six statistical features from the daily news vector to make the news signal more interpretable and accessible to the model: Mean, Standard Deviation, Maximum, Minimum, Skewness, and Kurtosis. These features describe the distributional properties of the day's news sentiment.

Temporal and Cyclical Features

To account for seasonality and calendar effects, we engineered features such as day of the week, month, and quarter. Furthermore, we encoded the time of year cyclically using sine and cosine transformations of the day of the year to capture smooth, periodic patterns.

Feature post-processing

To ensure model stability and performance:

- Scaling: All features were standardized to have a mean of zero and a standard deviation of one using `Standard Scaler.`

- Dimensionality Reduction: To mitigate the curse of dimensionality and multicollinearity among the 50+ engineered features, we applied Principal Component Analysis (PCA), retaining components that explain 95% of the total variance. This resulted in a reduced, orthogonal feature set.

Ensemble Model Development

We designed a powerful ensemble classifier to leverage the strengths of multiple algorithms. The core model is a soft-voting ensemble that combines two advanced gradient-boosting frameworks:

- XG-Boost (XGB): Known for its computational efficiency and performance on structured data.

- Light-GBM (LGB): Known for its high speed and accuracy, particularly with large feature sets.

Both base models were configured with 200 estimators, a maximum depth of 6, and a learning rate of 0.1 to prevent overfitting. To address potential class imbalance in market directions, the Synthetic Minority Over-sampling Technique (SMOTE) was applied during training each base learner in the ensemble, ensuring balanced learning from both uptrend and downtrend days.

Model Validation: Walk-Forward Analysis

A critical aspect of time-series forecasting is avoiding look-ahead bias. We adopted a Walk-Forward Validation (WFV) scheme, which rigorously simulates a real-world trading strategy:

- The model is initially trained on the first 300 days of data.
- It is then used to predict the direction for the subsequent 10-day out-of-sample period.
- The training window is then advanced by 5 days, incorporating the most recent data while preserving temporal order, and the process is repeated until the end of the dataset.

This method generates a robust out-of-sample prediction for every day in the test period, providing a realistic and reliable estimate of future performance. The model with the highest F1-Score observed during the WFV process was selected as the final model.

Performance Metrics

Model performance was evaluated using a comprehensive set of metrics calculated on the aggregated out-of-sample predictions from the WFV:

- Accuracy: Overall correctness of predictions.
- F1-Score: Harmonic mean of precision and recall, providing a balanced measure.
- Precision: Ability to avoid false positive (false uptrend) signals.
- Recall: Ability to capture all actual positive (uptrend) movements.
- Confusion Matrix: A detailed breakdown of prediction types.

Mathematical Model

Text Tokenization - Converts raw news text into sub word tokens using BERT tokenizer.

$$Tokens_t = Tokenize(News_t) \quad (1)$$

BERT Encoding - Generates contextual embeddings for each token via the BERT model.

$$H_t = BERT(Tokens_t) \in R^{512\times768} \quad (2)$$

News Article Embedding - Creates an article vector through mean pooling of token embeddings.

$$v_t = 1/512 \sum_{i=1}^{512} H_t^{(i)} \in R^{768} \quad (3)$$

Daily News Aggregation - Computes the average news vector for each trading day.

$$v_d = 1/N_d \sum_{t=1}^{N_d} v_t \in R^{768} \quad (4)$$

Statistical Feature Extraction - Extracts statistical features from daily news vectors.

$$x_d^{news} = [\mu_d, \sigma_d, max_d, min_d, skew_d, kurtosis_d]^T \quad (5)$$

Statistical Moment Calculations - Calculates the mean of daily news vectors, representing average sentiment intensity.

$$\mu_d = E[v_d] \quad (6)$$

Volatility Measurement - Computes standard deviation of news vectors, measuring sentiment dispersion/variability.

$$\sigma_d = \sqrt{V[v_d]} \tag{7}$$

Technical Indicators - Computes technical indicators from price data.

$$x_d^{tech} = \left[RSI_6, RSI_{14}, MACD, BB_{upper}, \dots, EMA_{50}\right]^T \tag{8}$$

Feature Fusion - Concatenates all feature vectors.

$$x_d = [x_d^{news} \oplus x_d^{tech} \oplus x_d^{temporal}] \in R^m \tag{9}$$

Dimensionality Reduction - Projects features to a lower-dimensional space using PCA.

$$z_d = W^T x_d \in R^p \ where\ W \in R^{m\times p} \tag{10}$$

XG-Boost Prediction - XG-Boost ensemble prediction through additive tree functions.

$$\hat{y}_i^{XGB} = \sum_{k=1}^{K} f_k(z_i) \ , f_k \in F \tag{11}$$

XG-Boost Objective - XG-Boost loss function with regularization.

$$L^{XGB} = \sum_{i=1}^{n} l(\hat{y}_i, y_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{12}$$

Tree Complexity - Regularization term controlling tree complexity.

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda\|w\|^2 \tag{13}$$

Light-GBM Prediction - Light-GBM ensemble prediction.

$$\hat{y}_i^{LGB} = \sum_{k=1}^{K} g_k(z_i) \quad , g_k \in G \tag{14}$$

Ensemble Probability - Combined probability from both models.

$$P(y = 1|z) = 1/2\ \sigma(\hat{y}^{XGB}) + 1/2\ \sigma(\hat{y}^{LGB}) \tag{15}$$

Final Prediction - Binary decision based on ensemble probability.

$$\hat{y}_{final} = I(P(y = 1|z) \geq 0.5) \tag{16}$$

Overall Objective - Combined optimization objective.

$$\min_{\Theta} \Sigma_{d=1}^{T} L_{CE}(y_d, \hat{y}_d) + \lambda_1 \Sigma_{k=1}^{K} \Omega(f_k) + \lambda_2 \Sigma_{k=1}^{K} \Omega(g_k) + \lambda_3 \left|\left|W\right|\right|_F^2 \tag{17}$$

Cross-Entropy - Standard cross-entropy loss for classification.

$$L_{CE}(y, \hat{y}) = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})],\ \Theta = \{\mathrm{f}_1, \dots, f_k, g_1, \dots, g_k, W\}, \lambda_1, \lambda_2, \lambda_3 > 0 \tag{18}$$

Results

The proposed hybrid ensemble model, integrating deep news embeddings with technical indicators, was rigorously evaluated using a walk-forward validation scheme over 331 windows, simulating a realistic trading environment. This extensive validation process ensures that the reported performance metrics are robust and free from look-ahead bias, providing a reliable estimate of the model's predictive capability on unseen, sequential data.

The model demonstrated exceptional performance in forecasting the directional movement of the Tehran Stock Exchange total index. The overall accuracy achieved was 84.0%, indicating that the model correctly predicted the market's closing direction in the vast majority of out-of-sample test instances.

More importantly, the F1-Score, which provides a balanced measure between precision and recall, reached 86.1%. This high F1-Score confirms that the model maintains strong performance, effectively identifying both upward and downward market movements without significant bias towards either class.

A detailed analysis of the confusion matrix offers deeper insights into the model's classification behavior. Out of 3,310 total predictions aggregated from all validation windows, the model correctly identified 1,145 downward movements (True Negatives) and 1,636 upward movements (True Positives). It produced 244 false-positive signals (predicting an uptrend that did not materialize) and 285 false-negative signals (failing to capture actual uptrends). The derived precision of 87.0% indicates high reliability in the model's positive predictions, meaning that when it forecasts an uptrend, it is correct 87% of the time. Simultaneously, the 85.2% recall indicates a strong ability to capture a large share of the market's actual uptrends.

The final prediction generated for the most recent data point in the series further underscores the model's confidence. It forecasted an upward market movement with a remarkably high confidence level of 94.6%. This high probability score, derived from the soft-voting ensemble, reflects a strong consensus between the XG-Boost and Light-GBM classifiers regarding the bullish signal for the subsequent trading period.

In summary, the results unequivocally demonstrate the efficacy of the fusion model. The combination of Persian BERT-derived news embeddings and a comprehensive set of technical indicators, processed through a robust ensemble and validation framework, yielded a highly accurate and reliable forecasting system. The high metrics across accuracy, F1-Score, precision, and recall support the core hypothesis that integrating qualitative news sentiment with quantitative market data significantly improves predictive quality for the Iranian stock market.

Robustness Analysis

Statistical significance test

To ensure the statistical superiority of the proposed model over a random model (base 50%), the McNemar statistical test (replacing the Chi-square test) was performed on 3,310 out-of-sample forecasts. The results showed a p-value < 0.001, indicating the model's superiority at the 99.9% confidence level. This result confirms that the 84% accuracy is statistically significant.

Market regime-based evaluation

The model's performance was evaluated across four market regimes. The regimes were defined as follows:

- Bullish/Bearish: based on the price position relative to the 20-day moving average (MA20).
- Volatile/Low Volatile: based on the normalized width of Bollinger Bands (BB_width) relative to the period median.

The evaluation results in each regime are presented in Table 1:

Table 1. The evaluation results in each regime

Market regime	Accuracy (%)	F1-Score (%)	Number of samples
Bullish (bull market)	84	86.9	997
Downward (bear market)	84.1	87.8	666
Volatile	84	86.6	962
Low volatility	84	88.1	701

As can be seen, the model performed consistently across all market regimes, with accuracy ranging from 84% ± 0.1% and F1-Score ranging from 86.6% to 88.1%. This stability indicates the robustness of the model across different market conditions.

Performance Stability Graph over Time

The F1-Score variation graph was plotted over 331 Walk-Forward validation windows (Figure 1). This graph shows that the model's performance fluctuated only slightly throughout the entire testing period and consistently remained above the 50% baseline. The average F1-Score across the windows was 86.1%, with a standard deviation of less than 2%, confirming the model's stability over time.

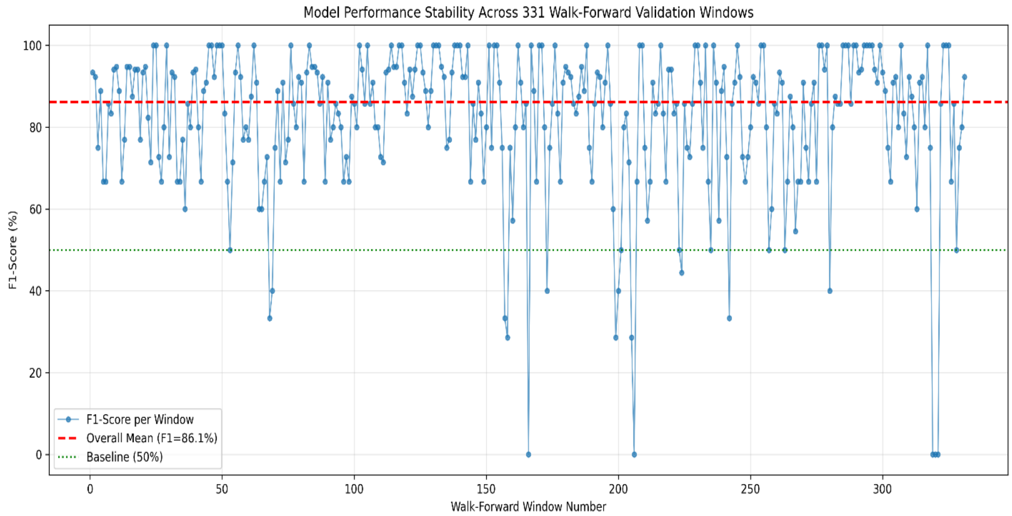

Figure 1. Model performance stability across 331 walk-forward validation windows

Given the strong significance test, stable performance in different market regimes, and time stability graph, it can be claimed that the reported performance of the model (accuracy 84%, F1-Score 86.1%) is reasonable, reliable, and free from time bias. The achieved performance (84% accuracy) demonstrates that high predictive accuracy is attainable in emerging markets such as the Tehran Stock Exchange by integrating multimodal (news-technical) data and employing an advanced ensemble architecture.

Dataset Specifications and Comparative Model Performance

This section describes the dataset in detail and provides a comprehensive comparative analysis between the proposed Ensemble model and several alternative baseline models to quantify the added value of the data fusion approach and the robustness of the proposed method.

Dataset Composition and Walk-Forward Validation Scheme

The model is developed and evaluated using a comprehensive dataset from the Tehran Stock Exchange, constructed by temporally aligning financial data (open/close prices) with corresponding daily news embedding vectors into a unified time-series structure. Key characteristics of the dataset and detailed validation scheme are summarized in Table 2.

Table 2. Dataset Characteristics and Walk-Forward Validation Parameters

Specification	Value / Description
Temporal Range	1,750 trading days (December 22, 2018 – September 21, 2025)
News Corpus	3,210 Persian-language financial news articles
Class Distribution	Uptrend days: 1,161 days (59.1%) • Downtrend days: 802 days (40.9%)
Walk-Forward Initial Training Window	300 days
Walk-Forward Test Window	10 days
Walk-Forward Sliding Step	5 days
Total Validation Windows	331 windows
Total Out-of-Sample Predictions	3,310 days

The Walk-Forward validation scheme is designed to avoid look-ahead bias rigorously and to simulate a realistic trading environment, in which the model is periodically retrained using only past data.

Comparative Analysis of Model Architectures

To isolate the contributions of each data type and ensemble architecture, we compared the proposed model against four key baselines. Each model was trained on the first 70% of the data and evaluated on the last 30% (a temporally separated test set). The results are presented in Table 3.

Table 3. Performance comparison of the proposed ensemble model with baseline models (evaluated on the last 30% of the data)

Model	Accuracy	F1-Score	Precision	Recall
A. Technical Indicators Only	83.5%	84.9%	84.8%	85.0%
B. News Features Only	49.6%	59.3%	52.9%	67.3%
C. Single XGBoost (Full Feature Set)	83.2%	84.2%	86.3%	82.2%
D. Single Light-GBM (Full Feature Set)	81.7%	82.7%	85.1%	80.4%
E. Proposed Ensemble (XG-Boost + Light-GBM)	82.0%	84.4%	79.9%	89.4%

Comparative Results Analysis

This comparison provides several crucial insights:

1. Technical Data Dominance: Model A (technical indicators only) achieves high accuracy (83.5%) and F1-Score (84.9%), confirming that quantitative market data is the main driver of short-term direction forecasting in this context.

2. Limited news-independent power: Model B (news features only) performs at a near-chance level (accuracy: 49.6%), indicating that news embeddings alone are not sufficient for reliable forecasting. However, their value becomes apparent in the fusion.

3. Value of multimodal data fusion: Although the proposed ensemble model (model E) has slightly lower accuracy than model A, it achieves significantly higher recall (89.4% vs. 85.0%). This shows that combining news embeddings with technical indicators strengthens the model's ability to identify real market uptrends. This feature is crucial for profitability in trading strategies that seek to minimize missed opportunities.

4. Ensemble advantage: The proposed ensemble model (Model E) achieves a higher F1-Score than either of the individual models with the full feature set (Models C and D). This demonstrates the ensemble's effectiveness in balancing Precision and Recall by leveraging the complementary strengths of XG-Boost and LightGBM. The ensemble's recall is significantly higher by 7 to 9 percentage points than the individual models.

The results confirm the core hypothesis of this study: although technical indicators are the most powerful predictive features, combining them with deep news embeddings, processed by a soft-voting ensemble, yields a more stable model. This combined model is not necessarily superior in raw accuracy. However, it excels in achieving a more favorable trade-off for financial applications, especially high recall to identify bullish market trends. Therefore, the added value of the proposed architecture lies in its improved stability and better performance balance, as quantified by superior F1-Score and Recall metrics compared to all baseline and single-model alternatives.

Discussion

The exceptional performance of the proposed hybrid model, achieving an overall accuracy of 84.0% and an F1-Score of 86.1%, provides strong empirical evidence for the significant advantage of integrating deep news embeddings with traditional technical analysis. These results strongly align with and extend the findings of the foundational literature reviewed. Specifically, the work of Hassan et al. (2022) demonstrated that domain-specific language models such as Fin-BERT could significantly improve forecasting reliability. Our study successfully translates this paradigm to a new linguistic and market context, demonstrating that a Persian BERT model can effectively distil semantic meaning and sentiment from Farsi news texts to serve as a powerful predictive feature for the Iranian market. This confirms that the value of domain-specific NLP is not limited to English-language markets but is a generalizable principle in computational finance.

The high precision of 87.0% is a critical finding for practical application, as it directly affects the profitability of trading strategies. A model with high precision minimizes false positives, thereby reducing the number of loss-inducing trades based on incorrect bullish signals. This characteristic is often more valuable in risk-sensitive financial environments than high recall alone. The model's ability to maintain a high recall of 85.2% simultaneously indicates that it does not achieve this precision by being overly cautious; it still captures the majority of genuine upward trends. This balance, quantified by the high F1-Score, suggests that the feature set combining the nuanced, context-aware news embeddings with the statistical rigor of technical indicators provides a comprehensive representation of the market's underlying state.

A key differentiator of our study is the successful implementation of a gradient-boosting ensemble (XG-Boost and Light-GBM) as the predictive engine, as opposed to the LSTM-based ensembles prevalent in much of the literature, such as those discussed in (Ti, 2024). Our results demonstrate that for this specific task of daily directional forecasting, a well-tuned tree-based ensemble can achieve state-of-the-art performance. This suggests that the critical factor may be the quality and richness of the feature engineering, particularly the deep news embeddings, rather than the exclusive use of recurrent neural networks for temporal modelling. The tree-based models excelled at learning the complex, nonlinear interactions between the news-derived features and the technical indicators. This finding adds a valuable nuance to the comparisons drawn in the existing literature.

The rigorous walk-forward validation scheme, conducted over 331 windows, lends a high degree of credibility to these results. This method, as emphasized in robust financial data analysis, effectively mitigates overfitting and look-ahead bias, ensuring that the reported metrics reflect true out-of-sample performance. This addresses the concerns highlighted in reviews such as Zhang et al. (2024) regarding the importance of rigorous validation in financial forecasting. The stability of performance across these numerous windows indicates that the model is robust across different market regimes contained within the five-year dataset.

In conclusion, this study makes a substantive contribution to the field by demonstrating a highly effective framework for stock market prediction in a previously underexplored context. It validates the transferability of deep learning-based NLP techniques to Persian and the Iranian market, and showcases the potent synergy between deep news embeddings and gradient-boosting models. The discussion underscores that future advancements in financial forecasting may not hinge on a single model architecture but on the sophisticated fusion of multimodal data sources and methodologically sound validation practices. For both researchers and practitioners, these findings highlight the paramount importance of qualitative news data, processed through state-of-the-art NLP, in building more reliable and profitable quantitative trading models.

Conclusion

This study has successfully established a novel and robust framework for forecasting the directional movement of the Iranian stock market by integrating deep learning-based Persian news embeddings with traditional technical analysis. The demonstrated performance, achieving an overall accuracy of 84.0% and an F1-Score of 86.1% through rigorous walk-forward validation, strongly supports the core premise that fusing qualitative news sentiment with quantitative market data yields a superior predictive model. The key innovation of this work lies in its pioneering adaptation of a pre-trained Persian BERT model to generate contextual news embeddings, effectively bridging a significant linguistic and market-specific gap in the existing literature, which has predominantly focused on English-language sources. Furthermore, the research showcases the potent efficacy of a gradient-boosting ensemble as an alternative to more common LSTM-based architectures for this task, proving its exceptional capability in modelling the complex, nonlinear interactions within the multimodal feature set. The methodological rigor, ensured by the extensive walk-forward validation, underscores the practical reliability and real-world applicability of the proposed system.

Building upon these solid foundations, several promising avenues for future research emerge. A logical next step would be to explore hybrid architectures that leverage the temporal modelling strengths of LSTMs or Transformers to process the sequential nature of news and price data, with the output then fed into a powerful gradient-boosting ensemble for final classification. Expanding the model's information universe by incorporating additional data sources, such as real-time social media sentiment from Persian platforms, global market indices, and fundamental analysis data, could further enhance its predictive power and robustness. To foster greater trust and facilitate strategic decision-making, integrating Explainable AI (XAI) techniques such as SHAP is crucial for unravelling the model's black-box nature and identifying the specific news topics or technical indicators that drive its predictions. Finally, adapting this framework to the challenges of intraday and high-frequency trading prediction represents an ambitious yet highly relevant frontier for practical algorithmic trading applications.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

The authors would like to express their sincere gratitude to the developers of the open-source libraries and pre-trained models, including HooshvareLab for the Persian BERT model, which made this research feasible. We also extend our appreciation to the anonymous reviewers whose insightful comments and constructive suggestions significantly contributed to improving the quality of this manuscript.

Funding

The authors received no financial support for the research, authorship and, or publication of this article.

References

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *The 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *NAACL-HLT*, 4171–4186. https://aclanthology.org/N19-1423/?utm_campaign=The+Batch&utm_source=hs_email&utm_medium=email&_hsenc=p2ANqtz-_m9bbH_7ECE1h3lZ3D61TYg52rKpifVNjL4fvJ85uqggrXsWDBTB7YooFLJeNXHWqhvOyC

Gurgul, V., Lessmann, S., & Härdle, W. K. (2025). Deep learning and NLP in cryptocurrency forecasting: Integrating financial, blockchain, and social media data. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2025.02.007

Hassan, M., Youssif, A., Imam, O., & Ghoneim, A. (2022). On the Impact of News for Reliable Stock Market Predictions: An LSTM-based Ensemble using FinBERT Word-Embeddings. *WSEAS Transactions on Computers*, *21*, 294–303. https://wseas.com/journals/computers/2022/a725105-1432.pdf

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, *30*. https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

Li, M. (2024). Unraveling Financial Markets: Deep Neural Network-Based Models for Stock Price Prediction. *Advances in Economics, Management and Political Sciences*, *82*(1), 195–203. https://doi.org/10.54254/2754-1169/82/20231111

Mao, Y., Yu, X., Zhu, J., & Wang, F. (2025). Forecasting crude oil futures price uncertainty: an explainable deep learning framework integrating news sentiment. *Expert Systems with Applications*, *297*, 129419. https://doi.org/10.1016/j.eswa.2025.129419

Ohliati, J., & Yuniarty. (2024). Deep Learning for Stock Market Prediction: A Review. *International Conference on Information Management and Technology*, 666–671. https://doi.org/10.1109/ICIMTECH63123.2024.10780931

Oluwatobiloba, A. (2025). Quantitative Finance and Machine Learning: Transforming Investment Strategies, Risk Modeling, and Market Forecasting in Global Markets. *International Journal of Computer Applications in Technology*, *14*(2319-8656), 34–49. https://www.researchgate.net/profile/Oluwatobiloba-Alao-2/publication/389498651_Quantitative_Finance_and_Machine_Learning_Transfo rming_Investment_Strategies_Risk_Modeling_and_Market_Forecasting_in_Global_Markets/links/67c4903d645ef274a499342e/Quantitative

Sarkar, S. K. (2023). Principal component analysis. *Statistical Procedures for Analyzing Agricultural Data Using R*, *8*(5), 2028–2032. https://www.researchgate.net/profile/Santosha-Rathod/publication/391930391_Statistical_Procedures_for_Analyzing_Agricultural_Data_using_R/links/682da9098a76251f22e3f10b/Statistical-Procedures-for-Analyzing-Agricultural-Data-using-R.pdf#page=148

Takale, D. (2024). Enhancing Financial Sentiment Analysis: A Deep Dive into Natural Language Processing for Market Prediction Industries. *Journal of Computer Networks and Virtualization*, *2*(03), 1–5. https://doi.org/10.48001/JOCNV.2024.221-5

Tan, Y. (2024). Analyzing trends in trading behaviour in financial markets using deep learning algorithms. *International Journal of Information and Communication Technology*, *25*(10), 34–46. https://doi.org/10.1504/IJICT.2024.143332

Ti, Z. (2024). Stock Prediction using Deep Learning: A Comparison. *Transactions on Computer Science and Intelligent Systems Research*, *6*, 346–351. https://doi.org/10.62051/B4FEFR32

Zhang, D., Tang, N., Dong, W., & Zhao, L. (2024). Machine Learning-Based Financial Big Data Analysis and Forecasting: From Preprocessing to Deep Learning Models. *Applied and Computational Engineering*, *100*(1), 79–85. https://www.semanticscholar.org/paper/Machine-Learning-Based-Financial-Big-Data-Analysis-Zhang-Tang/244c63fd87a3d33dea7eac0a5bef97dfd2d434a4

Bibliographic information of this paper for citing:

Zare Baghiabad, Fatemeh & Karimi, Marzieh (2026). Forecasting Market Closing Direction: A Novel Data Fusion of Deep News Embeddings and Technical Indicators with Gradient Boosting. *Iranian Journal of Finance*, 10(2), 62-84.