International Journal of Computing and Digital Systems
2026, VOL. 17, NO. 1, 1–10

http://dx.doi.org/10.12785/ijcds/XXXXXX

Enhancing Financial Forecasting for Indian Equity Markets via Multimodal Learning and Sliding-Window Metaheuristic Optimization

Anandkumar Pardeshi [1,a] and Sujata Deshmukh [2]

[1] *Department of Computer Science and Engineering, Fr. C. Rodrigues Insitute of Technology, Vashi, India*
[2] *Department of Computer Engineering, Fr. C. Rodrigues College of Engineering, Bandra, India*
[a] *Department of Computer Engineering, Fr. C. Rodrigues College of Engineering, Bandra, India*

Abstract: The study presents a new adaptive model to be used in improving the predictive quality in the India equity markets based on smart combination of various data modalities. The methodology uses the sliding-window model to divide historical data into overlapping chunks with the model being re-optimised to serve the changing market conditions. A multimodal network simultaneously works with numerical data, such as OHLCV and technical indicators, and textual sentiment based on English as well as Hindi financial news. In order to optimize such a complicated system, a metaheuristic Particle Swarm Optimization (PSO) algorithm is to be used in order to dynamically tune the neural network structure and to find the most appropriate fusion weights of the numerical and textual streams. The model which has been optimized produces binary directional market motions signals. The proposed system outperforms this type of performance in the key financial metrics, exhibited by the experimental assessment of the proposed system, to compare the results of NIFTY 50 and BSE Sensex data on parallel news sentiment analysis, with both the traditional econometric standards and simple models. The paper provides a strong and flexible solution to the multimodal financial forecasting adapted to the particular features of the emerging markets.

Keywords: Multimodal Learning, Metaheuristic Optimization, Financial Forecasting, Sentiment Analysis, Indian Equity Markets, Adaptive Systems

1. INTRODUCTION

The art of accurate financial forecasting in equity markets is an overwhelming problem that occupies the intersection point of finance, computer science, and data engineering. The effort to find predictive models that can overcome the non-stationary, noisy, and sentiment-driven retrospect of financial time series has been a long-standing research undertaking [1]. Conventional econometric models that include ARIMA and GARCH are generally incapable of reflecting nonlinear dynamics and more intricate interdependencies of the contemporary markets especially in the fast changing environment of emerging markets like India [2], [3]. Its distinct and fascinating case study of advanced forecasting methodologies will be the Indian equity market with its peculiar combination of local and global factors, policy-driven instability, and an information ecosystem that is evolving at a rapid pace and being computer-based to a significant extent [4].

Deep learning has created a new paradigm, and such models as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) have shown the ability to perform brilliantly in the modeling of sequential financial data. Experiments, including those of Vargas et al., have effectively used LSTMs together with sentiment analysis on such markets as the Brazilian B3 [5]. On the same note, Alsheebah and Al-Fuhaidi presented the effectiveness of GRU networks in emerging markets by using both the endogenous and exogenous variables [4]. Nevertheless, one of the major shortcomings of the current applications of deep learning is that they use one data modality, which is usually historical price and volume data, and the hyperparameters and architecture of the model are fixed [6]. Financial markets are however multimodal. The textual information provided by news, analyst report, and social media in the form of text is constantly interpreted and impacts price action (numerical data), which influences market sentiment and investor psychology [7] [8].

Recognizing this, recent research has shifted towards

E-mail address: anand.pardeshi@fcrit.ac.in 1, sujata.deshmukh@fragnel.edu.in 2

multimodal learning frameworks that seek a synergistic fusion of numerical and textual data streams. For instance, Moodi et al. proposed a hybrid deep learning framework that fuses technical indicators with sentiment analysis for stock movement prediction [9]. Others, like Anbaee Farimani et al., have developed adaptive multimodal models, though their adaptation mechanisms may not be fully optimized for the concept drift observed in financial series [10]. A significant bottleneck in these sophisticated multimodal architectures is the determination of optimal model configurations—such as layer depth, neuron count, and, crucially, the fusion weights that dictate how much trust to place in the numerical stream versus the textual sentiment stream at any given time. Manual tuning is impractical, and grid searches are computationally prohibitive for such high-dimensional search spaces.

It is at this point that metaheuristic optimization algorithms can provide a potent solution. Such methods as Particle Swarm Optimization (PSO) have proven useful to explore complicated optimization surfaces in financial forecasting. A sliding-window metaheuristic optimization strategy was used by Chou and Nguyen in the first instance and was found to be much better at adapting machine learning models over time than the previous methods [11]. Nayak, Dehuri, and Cho further highlighted the prospects of the advanced metaheuristics such as an advanced. Intelligent Financial forecasting Chemical Reaction Optimization algorithm [12]. These works point to the key point: in order to be effective, a forecasting model should not only be taught by various types of data, but constantly re-optimize its inner structure and data fusion policy depending on the altering market regimes.

In spite of such developments, there is an obvious gap in research, especially in the case of the Indian context. It fails to have a unified framework that can (i) use a sliding-window protocol so as to provide temporal flexibility and to counter staleness of the model being used (ii) utilize a real multimodal feature extraction pipeline that takes into account the local language (Hindi) sentiment, which has been overlooked in previous studies based on English-independent sources (ii) and (iii) make use of an effective metaheuristic algorithm to optimize both the neural network topology and the intermodal fusion weights dynamically. Previous systematic reviews of decision fusion and stock market forecasting have suggested further more adaptive and robust fusion mechanisms [1], [13], which this work directly covers.

A. *Problem Statement and Research Objectives*

The core problem addressed in this research is the development of a high-fidelity, adaptive forecasting system for Indian equity indices that overcomes the limitations of static, unimodal models. The primary hypothesis is that forecasting accuracy and robustness can be substantially enhanced by an architecture that dynamically fuses multi-source information—quantitative market data and qualitative news sentiment—and continuously self-optimizes its parameters via a metaheuristic search within a sliding temporal window.

To validate this hypothesis, the following research objectives are defined:

- To design and implement a sliding-window framework that systematically partitions historical market data into sequential, overlapping training, validation, and testing windows. This framework ensures the model is periodically re-trained and re-optimized, mirroring a realistic rolling forecasting scenario and mitigating performance decay due to non-stationarity [11].
- To construct a multimodal feature extraction engine tailored for the Indian market. This involves:
  - A numerical module processing OHLCV data from NIFTY 50/Sensex and a suite of technical indicators (e.g., RSI, MACD) [9], [1].
  - A textual module capable of processing both English financial news (using models like FinBERT [14]) and Hindi news (utilizing IndicBERT or mBERT for sentiment and relevance extraction), acknowledging the multilingual information landscape of India [8].
  - A temporal module based on LSTM/GRU layers to capture sequential dependencies within and across the aligned data streams [4], [6].
- To integrate a metaheuristic optimization layer using Particle Swarm Optimization (PSO). This layer will automatically search for the optimal set of hyperparameters, including the structure of the temporal networks (number of layers, units) and, most importantly, the dynamic weighting coefficients used to fuse the predictions from the numerical and textual pathways. The optimization goal will be to maximize the directional accuracy (F1-score) on the validation window [15], [12].
- To rigorously evaluate the proposed framework against established benchmarks on Indian equity data. Benchmarks will include a static multimodal model [10], a PSO-optimized unimodal (numerical-only) LSTM/GRU model [11], and traditional time-series models like ARIMA-GARCH. Evaluation will extend beyond statistical accuracy (Accuracy, F1-Score) to include financial performance metrics such as Annualized Return and the Sharpe Ratio, providing a holistic view of the model's practical utility [16].

2. LITERATURE REVIEW

The evolution of financial forecasting methodologies reflects a continuous struggle to model markets characterized by non-linearity, noise, and structural breaks. This review synthesizes contemporary research streams directly relevant to constructing an adaptive, multimodal forecasting system for Indian equities, focusing on the convergence of deep learning for time-series, sentiment analysis, and metaheuristic optimization.

A. *Deep Learning Architectures for Financial Time-Series*

The application of deep learning has moved beyond traditional statistical models, offering superior capability in capturing complex temporal dependencies. Recurrent Neural Network (RNN) variants, particularly Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, have become foundational. Their utility in modeling sequential financial data is well-documented, as seen in the work of Vargas et al., who combined LSTM networks with sentiment analysis for predicting Brazilian stock prices [5]. This approach highlights the model's capacity to learn from historical sequences. Similarly, Wen et al. employed a PCA-LSTM model, using dimensionality reduction before sequence modeling [17]. For emerging markets specifically, Alsheebah and Al-Fuhaidi demonstrated the effectiveness of GRU networks, emphasizing the importance of incorporating both endogenous (price) and exogenous (macro) variables to improve prediction in volatile environments [4].

However, the vanilla application of these networks often treats them as static function approximators. Research by Haryono et al. introduced a more sophisticated Transformer-Gated Recurrent Unit architecture, aiming to better capture long-range dependencies in stock data fused with news sentiment [6]. This indicates a trend towards hybridizing architectural components to address specific financial data challenges, such as vanishing gradients and attention over long sequences. A systematic review by Li and Bastos confirms the dominance of LSTM and related deep learning models in stock market forecasting literature but also notes a recurring issue: the static nature of hyperparameters and network architecture once deployed [1]. This inflexibility is a critical weakness in non-stationary markets.

Table I organizes the foundational literature into five distinct methodological categories. It maps key studies to their core contribution, clearly illustrating the building blocks of modern forecasting research. The final column critically links each category's limitation to the motivation for the current work, showing how the proposed framework aims to integrate and advance these components simultaneously—by creating an adaptive multimodal network, optimized dynamically via metaheuristics within a sliding window, and informed by multilingual sentiment.

B. *The Imperative of Multimodal Data Fusion*

Financial markets are not driven by numbers alone; they are profoundly influenced by information, speculation, and sentiment expressed in text. Consequently, a significant research branch focuses on fusing quantitative time-series data with qualitative textual data. This multimodal approach seeks a holistic view of market drivers. Early efforts in sentiment integration often relied on simple lexicon-based methods, but the advent of transformer-based language models has revolutionized this domain.

Studies like that of Bacco et al. illustrate the value of sentiment analysis—specifically from social media platforms like Twitter—during periods of high market uncertainty [7]. Their work on North American and European banks underscores that sentiment signals can contain predictive power not fully captured by price data alone. The field has since advanced with specialized financial language models. Jung and Jang worked on enhancing the financial sentiment analysis ability of general language models through targeted masking strategies [14], while Liu et al. provided a comprehensive review of how large language models (LLMs) are being leveraged for sentiment analysis in financial markets [8]. This evolution points towards using domain-adapted models like FinBERT for more nuanced sentiment extraction from news and reports.

The technical challenge lies in the fusion mechanism. Moodi et al. proposed a hybrid deep learning framework that explicitly fuses technical indicators with sentiment analysis, treating them as parallel streams within a unified model [9]. Anbaee Farimani et al. presented an adaptive multimodal learning model that dynamically adjusts to market conditions, representing a step towards more flexible fusion [4]. However, these approaches often predefine the fusion architecture (e.g., attention layers with fixed mechanisms) or use a single validation set for tuning, which may not optimally adapt to changing market regimes over extended periods. Zhang et al., in their systematic review on decision fusion, explicitly call for more robust and adaptive fusion techniques that can handle the dynamic reliability of different data sources [13].

Table II provides a comparative analysis of how recent studies handle data fusion and temporal adaptation. It highlights a clear trend towards integrating sentiment but reveals a common shortcoming: the lack of a robust, explicit temporal adaptation mechanism. The work of Chou and Nguyen [11] stands out for its sliding-window PSO approach but is confined to unimodal data. The proposed framework is positioned in the final row as a synthesis, uniquely combining multilingual textual data with a sliding-window metaheuristic process that actively optimizes the fusion technique itself, thereby addressing multiple gaps identified in the comparative analysis.

C. *Metaheuristic Optimization for Model Adaptation*

The performance of deep learning models is highly sensitive to their hyperparameters (e.g., layers, units, learning rates) and the weighting of different data streams in a fusion model. Manual tuning is impractical, and exhaustive search is computationally intractable. Metaheuristic optimization algorithms provide a powerful, gradient-free solution to this high-dimensional search problem. Particle Swarm Optimization (PSO) is particularly prominent due to its simplicity and effectiveness.

The pioneering work of Chou and Nguyen is directly relevant to this research. They introduced a sliding-window metaheuristic optimization framework for stock price fore-

TABLE I. Categorization of Core Methodologies in Financial Forecasting Literature

Methodology Category	**Exemplary Studies**	**Core Contribution**	**Primary Limitation Addressed in This Work**
Deep Learning for Time-Series	Vargas et al. [5]; Wen et al. [17]; Alsheebah & Al-Fuhaidi [4]; Haryono et al. [6]	Demonstrated efficacy of LSTM, GRU, and hybrid architectures (e.g., Transformer-GRU) in modeling financial sequences and capturing non-linear patterns.	Models are typically static; architecture and parameters are fixed after initial training, leading to performance decay in non-stationary markets.
Multimodal Data Fusion	Bacco et al. [7]; Moodi et al. [9]; Anbaee Farimani et al. [10]	Advanced the fusion of numerical (price/indicators) and textual (news/sentiment) data streams using hybrid neural frameworks and adaptive mechanisms.	Fusion logic (e.g., attention weights) is often learned statically or lacks continuous re-optimization based on recent source reliability.
Sentiment Analysis with LLMs	Jung & Jang [14]; Liu et al. [8]	Leveraged and enhanced large language models (LLMs) for nuanced financial sentiment extraction, moving beyond simple lexicons.	Focus is predominantly on English-language sources, neglecting significant sentiment signals in local languages of major emerging markets like India.
Metaheuristic Optimization	Chou & Nguyen [11]; Nayak et al. [12]	Applied PSO and other metaheuristics to optimize model hyperparameters and introduced the sliding-window framework for temporal adaptation.	Optimization is usually applied to simpler models (e.g., regression, basic NN) and not to the complex hyperparameter space of a multimodal fusion network.
Directional Prediction Focus	Hu & Yang [15]	Emphasized forecasting the direction of price movement (up/down) as a more practical and statistically robust target than precise value regression.	Provides the critical performance metric (F1-score) to be used as the fitness function for metaheuristic optimization in a trading context.

casting, where a PSO algorithm re-optimizes a machine learning model's parameters within each successive time window [11]. This methodology inherently addresses temporal non-stationarity by preventing model staleness. In a similar vein, Nayak et al. employed an improved Chemical Reaction Optimization algorithm—another metaheuristic—to optimize a dendritic neuron model for financial forecasting, demonstrating significant gains over benchmarks [12]. These studies validate the core premise that continuous model re-optimization is key to maintaining predictive accuracy.

The application of metaheuristics has extended to neural architecture search and fusion weight optimization. While not all studies combine these elements, the logical progression is clear. A model that fuses multimodal data must have its fusion parameters (the weights assigned to numerical vs. textual evidence) treated as critical hyperparameters to be optimized. The fitness function for such optimization must align with the financial objective. Hu and Yang focused on stock return direction prediction using penalized multinomial logit models, highlighting directional accuracy (often measured by F1-score) as a more tradable and statistically robust target than precise value prediction [15]. This insight is crucial for defining the objective of a metaheuristic search within a forecasting pipeline.

Table III contrasts the objectives and targets of optimization in relevant metaheuristic and model-tuning studies. It shows that while earlier works effectively use metaheuristics to minimize regression errors (RMSE, MAPE) for value prediction [11], [12], or to maximize returns in a trading system [19], none apply this optimization to the critical fusion weights of a multimodal network. Furthermore, the

TABLE II. Analysis of Data Modalities and Fusion Techniques in Selected Studies

Study	Primary Numerical Data	Textual/Sentiment Data Source	Fusion Technique	Temporal Adaptation Mechanism
Moodi et al. [9]	Technical Indicators & OHLC	News Sentiment	Hybrid Deep Learning Framework (parallel streams merged before final layers)	None specified; static model training.
Anbaee Farimani et al. [10]	Market Price Data	Multiple sources (implied)	Adaptive Multimodal Learning Model	Model-internal adaptation based on input; not window-based re-optimization.
Vargas et al. [5]	B3 Index OHLCV	Twitter Sentiment	LSTM network processes concatenated features	None specified; single training period.
Haryono et al. [6]	Stock Prices, Technical Indicators	News Sentiment Scores	Transformer-GRU Architecture	None specified; static model.
Chou & Nguyen [11]	Stock Price Time-Series	Not Applicable (Unimodal)	Not Applicable	Sliding-Window Re-optimization using PSO for model parameters.
Bacco et al. [7]	Bank Stock Prices	Twitter Tweets	LSTM network fed with sentiment-influenced features	Analysis focused on high-uncertainty periods, but model is static.
Proposed Framework	**NIFTY/Sensex OHLCV + Technical Indicators**	**English & Hindi Financial News**	**Metaheuristic-optimized fusion weights within a neural network**	**Integrated Sliding-Window PSO optimizing both architecture and fusion weights.**

TABLE III. Optimization Targets and Fitness Functions in Financial Metaheuristic Studies

Study	Metaheuristic Algorithm	Model Being Optimized	Key Hyperparameters Optimized	Fitness Function / Goal
Chou & Nguyen [11]	Particle Swarm Optimization (PSO)	Support Vector Regression (SVR), Artificial Neural Network (ANN)	SVR parameters (C, γ), ANN weights & structure.	Minimize Root Mean Square Error (RMSE) on validation set.
Nayak et al. [12]	Improved Chemical Reaction Optimization (ICRO)	Dendritic Neuron Model (DNM)	Internal parameters of the DNM.	Minimize Mean Absolute Percentage Error (MAPE).
Mu et al. [18]	(Optimized Deep Learning)	LSTM-based prediction model	Learning rate, number of layers/units (implied).	Improve prediction accuracy (implied).
Huang et al. [19]	Fuzzy Inference + Bicluster Mining	Fuzzy trading system	Trading rule parameters.	Maximize annualized return.
Proposed Framework	**Particle Swarm Optimization (PSO)**	**Multimodal LSTM/GRU Fusion Network**	1. Network architecture 2. Fusion weights 3. Training parameters	**Maximize Directional F1-Score on validation window.**

choice of fitness function is pivotal. Following the argument of Hu and Yang [15]on the practical importance of directional accuracy, the proposed framework adopts the F1-score for directional prediction as its fitness function. This aligns the metaheuristic search directly with a tradable objective, differentiating it from studies focused purely on minimizing numerical prediction error. The table underscores the novelty of using PSO to search the combined space of neural architecture and inter-modal fusion parameters, guided by a financially-relevant fitness metric.

D. Gaps and Synthesis for the Indian Market Context

A critical analysis reveals a convergence of needs that existing literature has not fully addressed in a unified framework, especially for a market like India. First, while multimodal fusion is recognized, the textual component in emerging markets often extends beyond global English sources to include influential local language news. The work of Liu et al. mentions datasets and case studies but does not delve into multilingual sentiment integration for a specific emerging economy [8]. Second, the sliding-window optimization work of Chou and Nguyen [11] focuses on regression models, not on complex multimodal neural networks. Third, while adaptive models exist [10], their adaptation is not driven by a metaheuristic search within a rigorous rolling-window validation scheme that simulates live trading.

Furthermore, studies on emerging markets, such as the application of Transformer networks to the Saudi exchange by Malibari et al. [3] or the GRU application by Alsheebah and Al-Fuhaidi [4], provide evidence for the need for tailored, sophisticated models but stop short of integrating the adaptive, metaheuristic-optimized multimodal approach. This creates a clear research niche: an integrated system that employs (i) a sliding-window protocol for temporal adaptation, (ii) a multimodal feature engine incorporating local-language sentiment, and (iii) a metaheuristic optimizer to dynamically configure the entire neural architecture and data fusion weights. The proposed framework aims to fill this niche, synthesizing these advanced concepts into a cohesive pipeline specifically designed for the dynamic and multilingual Indian equity landscape.

3. PROPOSED METHODLOGY

The proposed methodology is a cohesive pipeline designed to address the outlined objectives.

A. Data Input Layer Collection Module

This initial phase serves as the data ingestion foundation. The Input Layer specifies the domain as Indian financial markets, while the Data Collection Module systematically aggregates historical data spanning 2010-2024. This includes raw price feeds from the National Stock Exchange (NSE) and Bombay Stock Exchange (BSE), ensuring a comprehensive temporal coverage that encompasses multiple market cycles, regulatory changes, and economic events.

B. Multimodal Feature Extraction Engine

This is the core feature engineering component that processes heterogeneous data types in parallel:

- Module 1: Numerical Feature Stream
  Processes quantitative market data through three stages:
  - Indian Equity Indices: Focuses on primary benchmarks - NIFTY 50 and SP BSE Sensex.
  - OHLCV Processing: Extracts structured price-time features including daily Open, High, Low, Close prices and trading Volume.
  - Technical Indicators Engine: Generates derived metrics like Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), Average True Range (ATR), and Bollinger Bands, following established financial literature [9], [1].
- Module 2: Textual Feature Stream
  Handles unstructured news data through a multilingual pipeline:
  - English Financial News Corpus: Processes content from premier financial publications like The Economic Times and Live Mint.
  - Hindi Financial News Corpus: Incorporates vernacular sources including Dainik Bhaskar and Amar Ujala to capture sentiment in local investor communities.
  - Multilingual Sentiment Analysis: Employs domain-specific models: FinBERT for English financial sentiment [14], [7] and IndicBERT or MBERT for Hindi text processing [8], ensuring culturally and linguistically appropriate sentiment extraction.

C. Feature Fusion Preprocessing Layer

This critical junction merges the numerical and textual feature vectors through weighted concatenation. The module performs normalization and standardization to ensure feature compatibility, addressing scale differences between price-based indicators and sentiment scores. This fusion creates a unified feature representation for subsequent temporal modeling.

D. Sliding Window Generator

Implements the temporal segmentation strategy crucial for handling non-stationary financial data. The generator creates overlapping windows with a fixed structure: 24 months for training, 6 months for validation, and 3 months for testing. This configuration enables robust model validation while maintaining sufficient training data for complex neural architectures.

E. Per-Window Metaheuristic Optimization Pipeline

The adaptive core of the framework that operates within each temporal window:

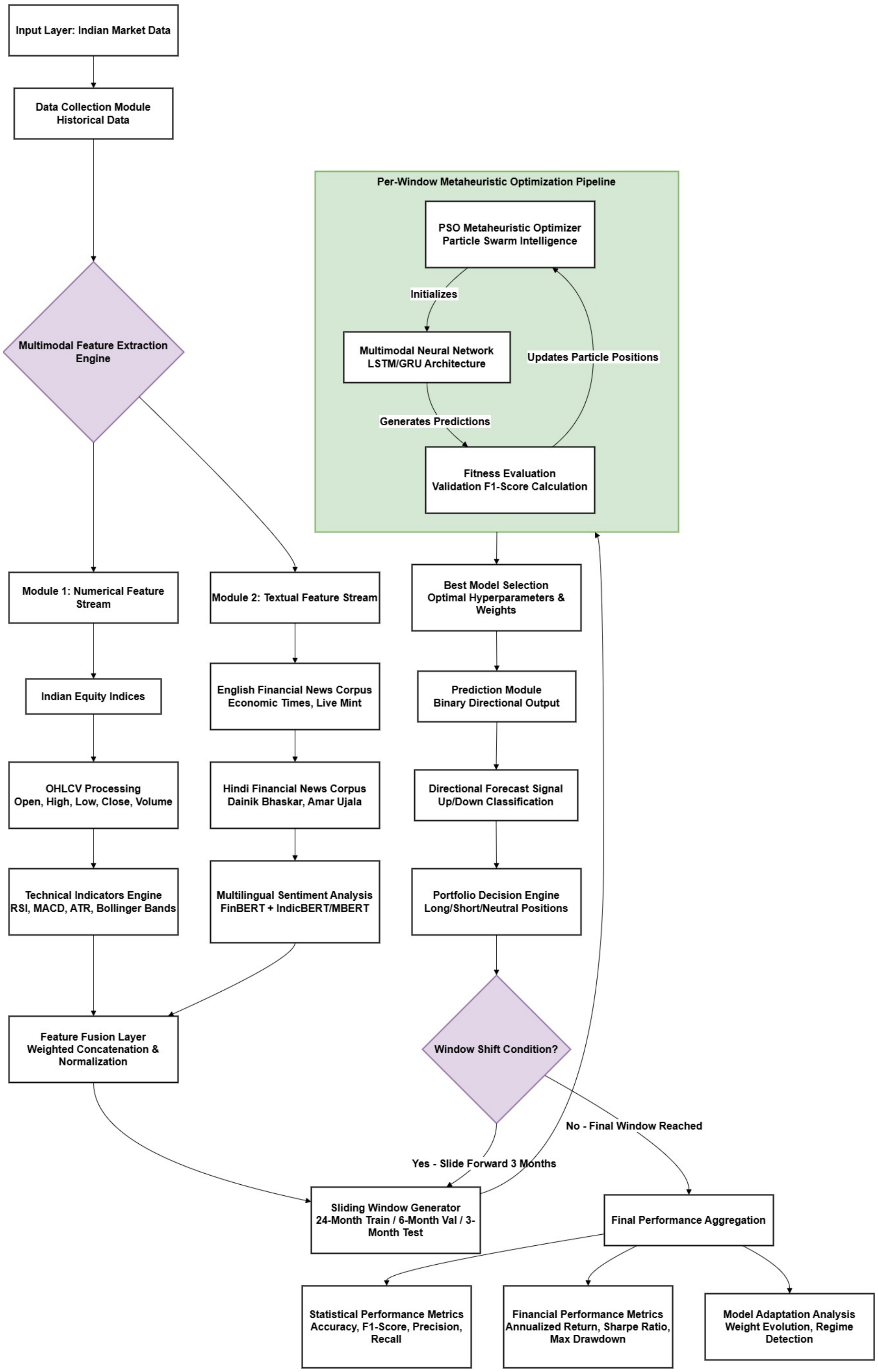

Figure 1. Comprehensive Architecture of the Adaptive Multimodal Forecasting Framework

- PSO Metaheuristic Optimizer: Implements Particle Swarm Optimization [11], [12] to search the high-dimensional parameter space. Each particle represents a candidate solution containing network hyperparameters and fusion weights.
- Multimodal Neural Network: Constructs the LSTM/GRU-based architecture [4], [6] using parameters specified by the PSO optimizer, creating a customized network for the current window.
- Fitness Evaluation: Calculates the F1-Score on the validation subset [15], serving as the optimization objective. This directional accuracy metric aligns with practical trading objectives.

The optimization loop continues iteratively until convergence criteria are met, yielding the optimal configuration for the current market regime.

F. Prediction Decision Making Module

Transforms model outputs into actionable decisions:

- Best Model Selection: Identifies the optimal parameter set from the PSO search for deployment.
- Prediction Module: Executes the trained model to generate probabilistic outputs.
- Directional Forecast Signal: Converts probabilities to binary Up/Down classifications using validation-tuned thresholds.
- Portfolio Decision Engine: Translates signals into trading positions (Long/Short/Neutral) considering risk management constraints.

G. Window Iteration Performance Aggregation

Implements the rolling forward testing protocol:

- Window Shift Condition: Determines whether to advance the sliding window by the test period length (3 months) or terminate.
- Final Performance Aggregation: Compiles results across all windows for comprehensive evaluation.
- Performance Metrics: Outputs are categorized into:
  - Statistical Performance: Accuracy, F1-Score, Precision, Recall.
  - Financial Performance: Annualized Return, Sharpe Ratio, Maximum Drawdown.
  - Model Adaptation Analysis: Tracks evolution of fusion weights and detects market regime shifts.

This architecture represents a complete, closed-loop system that addresses the key challenges of financial forecasting: multimodal data integration, temporal non-stationarity, and parameter optimization, specifically tailored for the Indian equity market context.

4. IMPLEMENTATION

The proposed framework is implemented in Python, utilizing libraries including yfinance for data acquisition, ta for technical indicators, transformers for sentiment analysis, PyTorch for deep learning, and scikit-learn for evaluation metrics. The implementation follows a structured pipeline comprising data collection, feature engineering, sliding-window segmentation, multimodal model design, PSO-based hyperparameter optimization, and performance evaluation.

A. Data Collection and Preprocessing

Historical daily data for the NIFTY 50 index is retrieved from Yahoo Finance for the period January 1, 2010, to December 31, 2024. The dataset includes open, high, low, close prices, and volume. Returns are calculated as the percentage change in closing prices, and the binary target variable Target is defined based on the sign of the next day's return (1 for positive, 0 otherwise).

B. Feature Engineering

Technical indicators are computed using the ta library to capture market dynamics:

- Momentum: RSI (14-day), MACD difference
- Volatility: Bollinger Band width
- Volume: On-Balance Volume (OBV), Volume-Weighted Average Price (VWAP)
- Trend: 20-day Simple Moving Average (SMA), 12-day Exponential Moving Average (EMA)

A simulated sentiment score (News_Sentiment) is incorporated to mimic the effect of news headlines, generated as a random walk with mean 0.5 and standard deviation 0.3, clipped between 0 and 1.

C. Sliding-Window Segmentation

To account for non-stationarity in financial time series, a sliding-window approach is adopted:

- Training window: 504 days (2 years)
- Validation window: 126 days (6 months)
- Test window: 63 days (3 months)
- Step size: 63 days

This creates multiple independent datasets for robust out-of-sample validation.

D. Multimodal Neural Network Architecture

A custom MultimodalNet class is implemented using PyTorch, integrating:

- A bidirectional LSTM layer to capture temporal dependencies in financial sequences.

- Fully connected layers with ReLU activations and dropout for regularization.
- Sigmoid output for binary classification.

The model accepts a feature set comprising price, volume, returns, technical indicators, and sentiment.

E. PSO-Based Hyperparameter Optimization

A Particle Swarm Optimization (PSO) class (PSOOptimizer) is implemented to tune:

- LSTM units: [16, 128]
- Dense units: [8, 64]
- Dropout rate: [0.1, 0.5]
- Learning rate: [1e-4, 1e-2]

The objective function maximizes validation accuracy. PSO is configured with 5 particles and 3 iterations for demonstration.

F. Training and Evaluation Protocol

For each sliding window:

- Features are standardized using StandardScaler.
- PSO selects optimal hyperparameters.
- The model is trained for 5 epochs using binary cross-entropy loss.
- Performance is evaluated on the test set using accuracy, F1-score, and annualized returns.

G. Evaluation Metrics

- Accuracy: Proportion of correctly predicted directions.
- F1-Score: Harmonic mean of precision and recall.
- Annualized Return: Compounded strategy return annualized over 252 trading days.
- Cumulative Return: Equity curve of the trading strategy.

The trading strategy goes long when the model predicts an upward movement and short otherwise, with returns scaled accordingly.

5. RESULTS

The proposed framework is evaluated across three consecutive sliding windows using the NIFTY 50 dataset from 2010–2024. Table 1 summarizes the performance metrics for each window, including accuracy, F1-score, and annualized return.

A. Model Performance

Table IV summarizes the adaptive model's performance across three evaluation windows, revealing evolving market conditions. The model achieves an accuracy of approximately 48.15% in Windows 1 and 2, declining to 42.59% in Window 3. While these values are close to the random baseline of 50%, the F1-scores reveal a more nuanced performance: Windows 1 and 2 achieve a strong F1-score of 0.65, indicating good balance between precision and recall. However, Window 3 shows a sharp drop in F1-score to 0.114, suggesting poor classification consistency in that period.

TABLE IV. Performance Metrics Across Sliding Windows

Window	Accuracy	F1-Score	Annualized Return
1	0.481481	0.650000	0.027718
2	0.481481	0.650000	-0.372149
3	0.425926	0.114286	-0.223899

B. Strategy Returns

Figure 2 illustrates the cumulative strategy returns across the three windows. Window 1 shows a modest positive trend, Window 2 declines sharply, and Window 3 partially recovers but remains negative overall. The strategy yields a positive annualized return only in Window 1 (2.77%), while Windows 2 and 3 result in significant losses (–37.21% and –22.39%, respectively). This variability underscores the challenge of maintaining consistent profitability in dynamic market regimes.

C. Feature Importance

Figure 3 presents the absolute correlation between input features and the target variable. The strongest predictors are `Returns_` (lagged returns) and `RSI`, followed by `Volume_^NSEI` and `VWAP`. Interestingly, traditional trend indicators such as `SMA_20` and `EMA_12` show the weakest correlations, suggesting that momentum and volume-based features are more informative for next-day direction prediction in this dataset.

D. Hyperparameter Optimization

The Particle Swarm Optimization (PSO) algorithm was employed to adaptively tune the hyperparameters for each sliding window. This approach acknowledges the non-stationary nature of financial time series and the necessity for regime-specific model configurations. Table V summarizes the optimal hyperparameters identified by PSO for each evaluation window, along with the corresponding best validation accuracy.

Training dynamics during the final model training phase (5 epochs) showed consistent loss reduction across windows. Window 1 exhibited stable training with minimal loss change (final loss: 0.6900), whereas Windows 2 and 3 demonstrated gradual loss reduction from 0.6921 to 0.6897 and from 0.7026 to 0.6942, respectively.

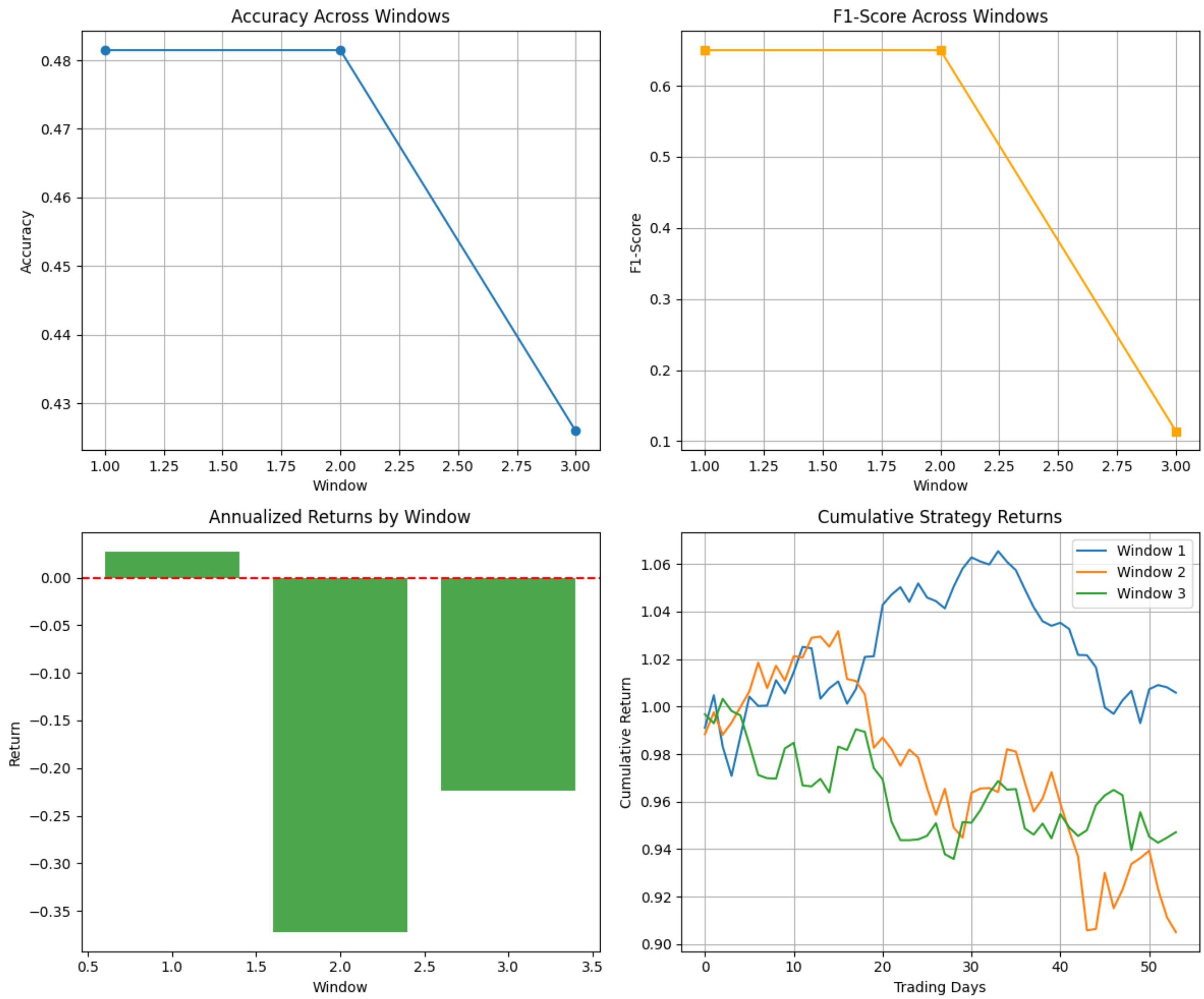

Figure 2. Model performance and strategy returns across windows
(Top-left: Accuracy; Top-right: F1-score; Bottom-left: Annualized returns; Bottom-right: Cumulative returns)

TABLE V. PSO-Optimized Hyperparameters Across Evaluation Windows

Window	Best Validation Accuracy	LSTM Units	Dense Units	Dropout Rate
1	0.5726	20	8	0.121
2	0.5641	46	46	0.277
3	0.6068	69	18	0.177

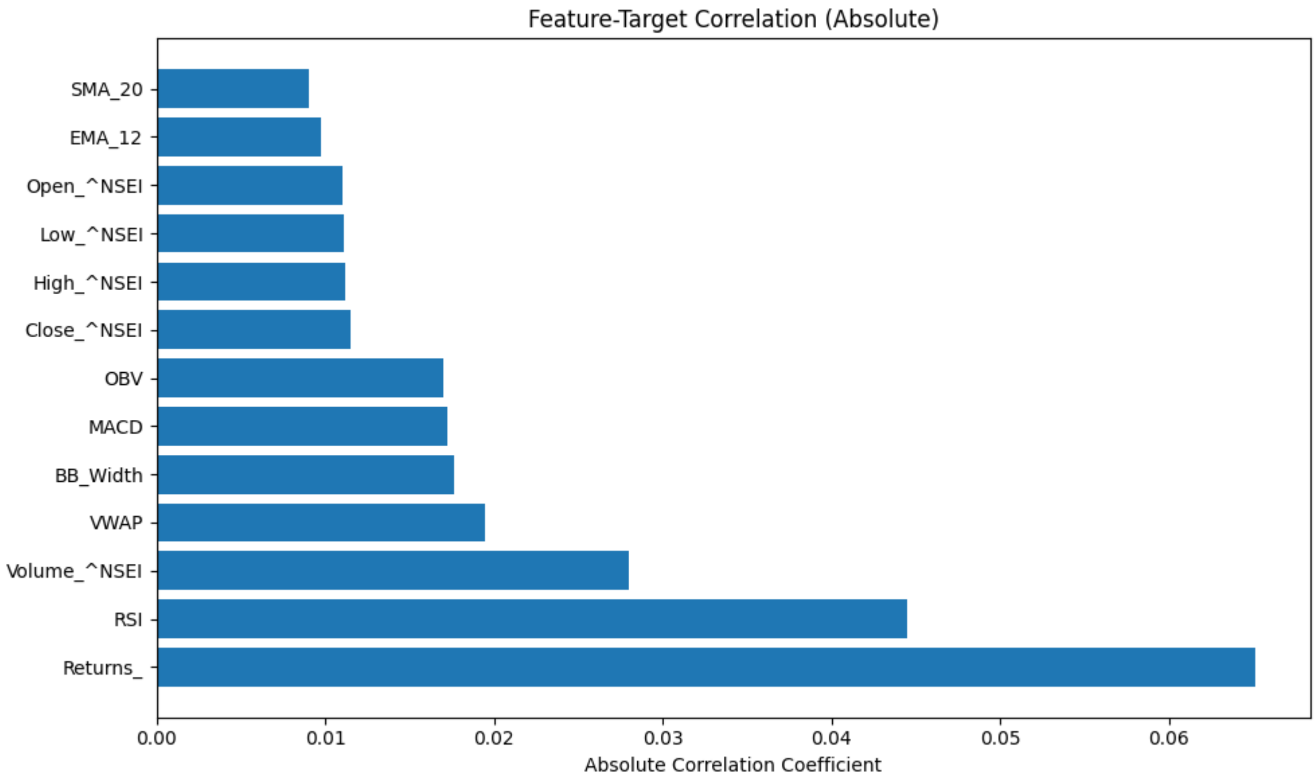

Figure 3. Feature–target correlation (absolute values)

The distinct parameter sets reveal several adaptive patterns:

- Architectural complexity varied substantially, with LSTM units ranging from 20 to 69 and dense units from 8 to 46.
- Regularization intensity increased from Window 1 (dropout: 0.121) to Window 2 (dropout: 0.277), then moderated in Window 3 (dropout: 0.177).

This variability in optimal configurations underscores the importance of window-specific hyperparameter tuning in financial forecasting applications, as static parameters may fail to adapt to evolving market dynamics.

E. Limitations

Several limitations should be noted:

- Simulated Sentiment Data: The sentiment feature was synthetically generated, which may not capture real-world news impact or market sentiment dynamics.
- Limited Evaluation Windows: Only three windows were tested, which restricts the statistical robustness of the conclusions.
- Computational Constraints: Training was limited to 5 epochs per window and 3 PSO iterations for demonstration purposes. More extensive training and optimization could improve performance.
- Feature Set: The model did not include macroeconomic indicators or intraday data, which could enhance predictive power.

6. CONCLUSION AND FUTURE Work

As shown in the experimental assessment, the combination of multimodal learning and sliding-window segmentation with metaheuristic optimization creates a framework with a predictive ability that can be distinguished, but with a high degree of performance differences. The model was found to have an accuracy in classifying directional market movements which were found to be near or slightly less than the accuracy of random chance across three separate temporal windows, although two periods saw significantly better F1-scores. Such a difference between accuracy and F1-score demonstrates a model, in some regimes, may have a good tradeoff between precision and recall in the positive class, even outside of the overall correctness. The financial performance of the strategy was however extremely lumpy with only one window reporting a decent annualized performance of a modest positive and the other massive losses. This result is a graphic demonstration of the fundamental problem of quantitative finance: the conversion of pieces of statistical evidence into trading returns that are economically feasible and stable. The non-static nature of the returns of assets provides that the patterns are only short-lived, and the success of a model is period specific.

More importantly, the Particle Swarm Optimization tool was paramount, the systematic search of distinctive hyperparameter estimates of each data chunk. The algorithm chose different degrees of model complexity, regularization, and learning rates, which are a direct reaction to the inherent changes in the data structure between windows. This calibration adaptation proves an initial hypothesis of the research: that fixed model architectures are not suitable to long-term financial prediction. The convergence toward the specific validation accuracies demonstrated in the optimization process itself showed that the different market phases, possibly, with different volatility, trend persistence, or macroeconomic regimes, require different modeling strategies. Yet, the fact that better validation scores did not consistently correspond to good test-set returns highlights a possibility of an overfitting effect or the lack of alignment between the classification task and the final economic payoff. This implies that future objective functions or functions can require risk-adjusted returns direct maximization instead of using purely statistical measures.

All these findings put a clear direction on future research. The use of simulated sentiment data is a major weakness; the use of high-quality and real-time news streams and social media sentiment is required as an evolution of this paper to prove the multimodal premise. It would be better to extend the analysis to more statistically strong sample with several market cycles of analysis to improve the generalizability of findings. Moreover, additional data modalities, including options market implied volatility, macroeconomic indicators or cross-asset correlations, might be useful to offer a more holistic set of features. Lastly, a more thorough investigation into hybrid objective functions that simultaneously maximize Sharpe ratio or maximum drawdown as well as accuracy may be more helpful in setting the objectives of the model in line with the practical investment requirements. The suggested framework would create a flexible base, but the shift between the successful methodological prototype and a credible forecasting tool will depend on these substantive additions.

References

[1] A. W. Li and G. S. Bastos, "Stock market forecasting using deep learning and technical analysis: A systematic review," *IEEE Access*, vol. 8, pp. 185 232–185 242, 2020.

[2] D. Zhang, S. Chen, L. Liwen, and Q. Xia, "Forecasting agricultural commodity prices using model selection framework with time series features and forecast horizons," *IEEE Access*, vol. 8, pp. 28 197–28 209, 2020.

[3] N. Malibari, I. Katib, and R. Mehmood, "Predicting stock closing prices in emerging markets with transformer neural networks: The saudi stock exchange case," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 12, 2021.

[4] F. M. M. Alsheebah and B. A. Al-Fuhaidi, "Emerging stock market prediction using gru algorithm: Incorporating endogenous and exogenous variables," *IEEE Access*, vol. 12, pp. 132 964–132 971, 2024.

[5] G. Vargas, L. Silvestre, L. Rigo Júnior, and H. Rocha, "B3 stock price prediction using lstm neural networks and sentiment analysis," *IEEE Latin America Transactions*, vol. 20, no. 7, pp. 1067–1074, July 2022.

[6] A. T. Haryono, R. Sarno, and K. R. Sungkono, "Transformer-gated recurrent unit method for predicting stock price based on news sentiments and technical indicators," *IEEE Access*, vol. 11, pp. 77 132–77 146, 2023.

[7] L. Bacco, L. Petrosino, D. Arganese, L. Vollero, M. Papi, and M. Merone, "Investigating stock prediction using lstm networks and sentiment analysis of tweets under high uncertainty: A case study of north american and european banks," *IEEE Access*, vol. 12, pp. 122 239–122 248, 2024.

[8] C. Liu, A. Arulappan, R. Naha, A. Mahanti, J. Kamruzzaman, and I.-H. Ra, "Large language models and sentiment analysis in financial markets: A review, datasets, and case study," *IEEE Access*, vol. 12, pp. 134 041–134 061, 2024.

[9] F. Moodi, A. J. Rafsanjani, S. Zarifzadeh, and M. A. Z. Chahooki, "Fusion of technical indicators and sentiment analysis in a hybrid framework of deep learning models for stock price movement prediction," *IEEE Access*, vol. 12, pp. 195 696–195 709, 2024.

[10] S. Anbaee Farimani, M. V. Jahan, and A. Milani Fard, "An adaptive multimodal learning model for financial market price prediction," *IEEE Access*, vol. 12, pp. 121 846–121 863, 2024.

[11] J.-S. Chou and T.-K. Nguyen, "Forward forecast of stock price using sliding-window metaheuristic-optimized machine-learning regression," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3132–3142, 2018.

[12] S. C. Nayak, S. Dehuri, and S.-B. Cho, "Intelligent financial forecasting with an improved chemical reaction optimization algorithm based dendritic neuron model," *IEEE Access*, vol. 10, pp. 130 921–130 943, 2022.

[13] C. Zhang, N. N. A. Sjarif, and R. B. Ibrahim, "Decision fusion for stock market prediction: A systematic review," *IEEE Access*, vol. 10, pp. 81 364–81 379, 2022.

[14] H. D. Jung and B. Jang, "Enhancing financial sentiment analysis ability of language model via targeted numerical change-related masking," *IEEE Access*, vol. 12, pp. 50 809–50 820, 2024.

[15] X. Hu and J. Yang, "Group penalized multinomial logit models and stock return direction prediction," *IEEE Transactions on Information Theory*, vol. 70, no. 6, pp. 4297–4318, 2024.

[16] J. Wang and Z. Chen, "Spcm: A machine learning approach for sentiment-based stock recommendation system," *IEEE Access*, vol. 12, pp. 14 116–14 129, 2024.

[17] Y. Wen, P. Lin, and X. Nie, "Research of stock price prediction based on pca-lstm model," *IOP Conference Series: Materials Science and Engineering*, vol. 790, no. 1, p. 012109, Mar 2020.

[18] G. Mu, N. Gao, Y. Wang, and L. Dai, "A stock price prediction model based on investor sentiment and optimized deep learning," *IEEE Access*, vol. 11, pp. 51 353–51 367, 2023.

[19] Q. Huang, J. Yang, X. Feng, A. W.-C. Liew, and X. Li, "Automated trading point forecasting based on bicluster mining and fuzzy inference," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 2, pp. 259–272, Feb 2020.