Wavelet-Based Dual-Domain Phase Alignment for Predicting Stock Index Trends from Investor Sentiment Cycles

Monika Gorkhe[1], Dr. Diksha Tripathi[2], Dr. Pravin D Sawant[3], Elangovan Muniyandy[4], Dr. Veera Ankalu Vuyyuru[5], Dr. Pratik Gite[6], Raman Kumar[7]
Department of School of BFSI, Symbiosis Skills and Professional University, Pune, India[1, 2]
Associate Professor, Department of Commerce, Narayan Zantye College of Commerce, Bicholim, Goa, India[3]
Department of Biosciences-Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India[4]
Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, 522502, A.P, India[5]
Associate Professor, Department of Computer Science and Engineering, Parul University, Parul Institute of Technology, Vadodara (G.J.), India[6]
University School of Mechanical Engineering, Rayat Bahra University, Mohali, India[7a].
Faculty of Engineering, Sohar University, Sohar, Oman[7b]

Abstract*—The dynamics in the financial markets are complicated and non-stationary, with a significant influence of the wave of investor sentiment. The recent changes in sentiment-based stock prediction have shown promising results, but the current research is to a large extent based on individual domain analysis, constant correlation, or the traditional machine learning framework, which constrains its capability to elucidate multi-scale temporal dynamics and phase-based lead-lag relationships. In order to overcome these weaknesses, a new Cross-wavelet Sentiment-driven Dual-domain Phase Alignment, abbreviated as CS-D²PA, is introduced for stock index trend prediction. The suggested structure has 91.8, 90.6, 92.1, 91.3, and 93.5 accuracy, precision, recall, F1-score, and trends consistency rate, respectively, proving to have a better predictive stability and a better classification performance in sentiment-driven stock trend forecasting. The non-stationary and multi-scale behavior of financial markets is dictated by non-periodic changes in investor sentiment cycles. This work proposes a Cross-wavelet Sentiment-based Dual-Domain Phase Alignment model (CS-D2PA) of predictive modeling of stock index trend. The framework combines the feature extraction by the continuous wavelet technique with the cross-wavelet phase difference estimation, as well as the structured alignment in time and frequency domains. The explicit modeling of the sentiment-price phase synchronization of the approach makes it possible to identify lead-lag interaction early and increases the predictability of the forecasts in volatile market conditions, which increases their interpretability.

Keywords*—*Stock trend prediction; investor sentiment analysis; wavelet time–frequency analysis; phase synchronization modeling; financial market forecasting

I. INTRODUCTION

Behavioral and psychological factors are increasingly gaining control over financial markets, thus investor sentiment is a key determinant of the dynamics of stock indexes [1],[2]. Conventional financial theories are based on rationality in decision making, but the use of emotionality, opinion, and group expectations has been found to greatly influence price changes, which are empirically supported [3],[4]. Increasingly, in the digital space, a great deal of sentiment-based information about the market has been created through news articles, social media posts and online forums, allowing the opportunity to analyze the market sentiment [5]. The recent developments of signal processing and financial analytics have aimed at identifying cyclicality in sentiment indicators to describe market changes [6]. Among other techniques, wavelet methods have attracted interest owing to the possibility that they enable the signals of non-stationary financial data to be decomposed into several time-frequency elements [7]. These techniques enable the researcher to examine both long- and short-term relationships between sentiment and the stock index [8],[9]. Nonetheless, the majority of the available methods are more concerned with correlation or coherence analysis, which does not provide much predictive power. Additionally, they tend to ignore the phase dependencies that indicate lead-lag relationships between the cycles of feelings and trends in the markets. Since financial markets are multi-scale time-dependent, predicting the fit between sentiment phases across domains or their deviations is crucial to proper prediction. This dynamic environment points to the importance of powerful structures that are able to connect cycles of sentiments and market indicators, maintaining temporal and frequency-related traits.

Although significant progress has been made, the recent predictive models have a number of methodological issues when used in the sentiment-based stock index forecasting [10]. The modern research is dominated by machine learning and deep learning models that are based on large datasets and various complex architectures [11]. Even though those models are quite accurate, they are frequently uninterpretable and do not model explicit cyclical coordination of sentiment and price movements. Besides, most studies work within one analytic space, either time or frequency, and thus, they are unable to investigate dynamic phase changes that exist in financial signals. The cross-wavelet and the wavelet coherence methods have been proposed to solve the nonlinearity and non-stationarity, but their use is more of a description than a prediction [12],[13]. The current frameworks hardly use phase alignment as a prediction mechanism, leading to delayed or noisy forecasts. More so, sentiment indicators are often assumed to be constant explanatory variables without taking into account their changing cyclical characteristics with time variation between time scales. Such limitations make the prevailing forecasting systems weak, particularly in turbulent market environments. To circumvent these limitations, the need to create dual-domain analytical frameworks, which would bring sentiment and price stages into harmony, is increasing. This approach can be used to detect trends at earlier stages, make predictions more reliable, and gain a better understanding of sentiment-based market processes.

A. Problem Statement

Even though investor sentiment has been greatly considered as a major factor that affects stock index movements, the available works lack critical limitations [14]. The majority of sentiment-based prediction models focus on the strength of correlation or classification performance without focusing on the phase-based lead-lag relationship that is important in trend forecasting [15]. Wavelet-based methods are typically limited to coherence analysis, where it can be described, but phase information cannot be incorporated in predictive decision-making [16]. Also, the effective number of frameworks is limited to one field without the combined use of time-domain and frequency-domain properties of financial signals [17]. Algorithms based on machine learning, though potentially powerful, are often not transparent, and do not cope with the non-stationary conditions of the market [18]. Moreover, sentiment cycles are generally studied in isolation without considering their synchronization with the stock index movements on various scales [19]. Such shortcomings restrict the applicability of the current models to explain early changes of direction and dynamic reactions of the market. This, therefore, presents a definite requirement for a predictive model that clearly interrelates sentiment and market lifecycle under two domains to enhance stock index trend forecasting.

B. Research Significance

The study presents a new cross-wavelet-based dual-domain phase aligning model that improves the prediction of stock index trends using sentiment cycles. The proposed solution enhances the interpretability, multi-scale market dynamics, and early trend signals, which are added by the implementation of phase synchronization into a predictive structure, generating more accurate and explainable financial forecasting models

C. Research Motivation

This study has been driven by the fact that sentiment analysis has always been at variance with the actionable market prediction. Existing models do not make use of cyclical phase behavior sufficiently, leading to late forecasts. This study can transform insights made by sentiments into eco-friendly, predictable, and scalable predictions in complex and volatile financial marketplaces by constructing the CS-D²PA framework.

D. Key Contributions

- Introduces a new and sophisticated sentiment-price interaction framework to address the shortcomings of traditional stock trend prediction methodologies.
- Takes advantage of synchronous time and frequency dynamics of sentiment to extract meaningful market behavior to look beyond short-term noise.
- Presents a cross-wavelet-based dual-domain phase alignment model that allows interpretable and effective financial time-series modeling.
- Multi-scale phase synchronization of sentiment and price signals, which enhances consistency in volatilities and stabilized market conditions.
- Achieves 91.8% accuracy, 93.5% trend consistency, which validates strong and consistent sentiment-driven stock trend predictions.

E. Rest of the Sections

In section II, prior methods of sentiment-driven stock prediction and wavelet-based analytical methods are discussed, and their limitations to capture phase synchronization and multi-scale dependencies are mentioned. Section III gives the CS-D²PA framework comprising sentiment processing, cross-wavelet analysis and phase alignment of the two domains. Section IV reports experimental analysis, performance analysis and predictive stability analysis. The findings, contributions, and future improvements to strong financial forecasting are concluded in Section V.

II. Related Works

Cristescu et al., [20] studied the impacts of the public sentiment, which is retrieved through financial news headlines and descriptions, on the stock prices of large technology companies. The main aim was to estimate the extent to which the sentiment polarity can explain the price changes in Microsoft, Tesla, and Apple stocks. The authors used Pearson correlation, regression analysis and wavelet coherence to achieve the use of both time and frequency relationships. The collected data were news data and daily stock prices. Their findings were able to confirm that sentiment polarity is an important factor in certain stock price changes, and that it has lagged effects across time scales. Nevertheless, the research is descriptive in character and fails to incorporate the information of phases within a predictive model. Also, the emphasis on large-cap companies makes it less generalizable, and the lack of a forecasting mechanism makes it less applicable to real-time market prediction.

Mamillapalli et al. [21] concentrated on enhancing the predictability of the stock price with sentiment indicators incorporated in the machine learning models. The objective of the research was to determine the performance of sentiment-enhanced models compared to conventional prediction methods when the market is volatile. The features of sentiment were extracted with the help of a lexicon-based sentiment analysis approach and were integrated into a gated recurrent unit model called GRUvader. The data used were historical stock price data and textual sentiment data. The results showed that AI-based models are far better than standalone statistical methods. The model is not interpretable, and it does not explicitly analyze the cycles of sentiment over time, although there is increased prediction power. In addition, the framework can only work in the time domain without considering frequency-dependent sentiment behavior that is important in multi-scale financial forecasting.

Nguyen et al., [22] determined how investor sentiment affects the stock market returns in the Vietnamese market, both in the short and long term. The objective of the research was to identify relationships of equilibrium and predictive ability of sentiment indicators. The authors used Vector Error Correction Models, regression methods, and various machine learning algorithms to analyze Facebook-based sentiment data derived in 2013 and 2023 using VNIndex returns and existing 2013-2023 data. The findings revealed that sentiment is a more predictive factor than previous price data, particularly when the market is at its extremes. Although the study has been able to prove the predictive power of sentiment, it heavily depends on statistical and machine learning tools without considering the dynamics of non-stationary phases. Moreover, the interactions between sentiment and prices are not discussed at time-frequency scales, which hinders the initial trend identification.

Ren et al., [23] investigated the predictability of stock market returns to investor sentiment derived from social media. They aimed at comparing various methods of sentiment representation and evaluating them in terms of their forecasting ability. Data on Twitter posts in 2019 and 2021 were trained to construct various sentiment indices based on lexicon-based algorithms, embedding algorithms, and language model fine-tuning. Prediction was done using machine learning models like SVR, Random Forest and AdaBoost. The analysis revealed that composite sentiment indices are better than single sentiment predictors and provide better returns forecasting. But the method takes sentiment as a cyclical numerical input, ignoring the cyclical nature of the input. Its failure to provide time-frequency or phase-sensitive modeling limits it to dynamic sentiment-market-synchronization.

Zhang et al., [24] examined the effects of the investor sentiment on the returns of cryptocurrencies, especially during holiday seasons. The study sought to know the way sentiment mediates market anomalies like the holiday effect. The authors used the fixed-effects regression models based on cryptocurrency trading data from 2017-2022 and sentiment indicators in the textual analysis. Their results showed that positive sentiment is strengthening returns and reducing the holiday effect. The study, though insightful, is confined to inference regression and does not mention predictive forecasting. Moreover, the sentiment dynamics are modeled in an aggregated manner without having multi-scale time dependencies and, hence, the model is not applicable in high-frequency or early trend forecasting.

Zhao et al., [25] studied the correlation between investor feelings based on stock forum comments and index price changes of the CSI 300. The objective of the study was to measure sentiment and determine the relationship between sentiment and market behavior. Sentiment indicators based on textual data were built with the help of sentiment dictionaries, and the correlation analysis, along with deep learning techniques, was utilized. The findings indicated that there is a positive correlation between sentiment and stock prices, which is lagged through time. Nevertheless, the methodology involves more correlation as opposed to prediction. The phase- level interpretation is absent in the sentiment indicator, and a declining effect over time is not simulated across the frequency bands; it does not help predict the market turning points.

Sable et al., [26] provided a thorough overview of stock market-based sentiment analysis methods. The research attempted to make comparisons between lexicon-based, machine learning and hybrid models in terms of scalability, accuracy and computational complexity. There were 57 research articles that were published between 2019 and 2024 analyzed. To standardize the assessment of the models the authors proposed a Stock Sentiment Rank Metric. Although the work offers good methodological information, it fails to offer a predictive framework. Additionally, the sentiment dynamics are considered qualitatively, not modeled at the level of signals (temporally or frequency-based), and thus, an opening is available to phase-sensitive and predictive sentiment alignment methods

Deng et al., [27] developed a combined machine learning model to predict the direction of stock index movement based on investor sentiment. This was aimed at increasing the trading performance and prediction accuracy. The research used lightGBM, NSGA-II optimization, and sliding window methods by using sentiment information of individual, institutional and foreign investors and Shanghai Stock Exchange index data. The proposed model recorded high hit ratios and trading returns. Although there has been great performance, the method is based on black-box optimization, and is not interpretable. Further, sentiment is modeled as an input feature as opposed to a dynamic cyclical signal and no phase synchronization mechanism is built.

Table I presents a comparison of some of the recent studies that have been conducted on sentiment-driven financial analysis based on the methods, data, strengths and limitations used. Although earlier studies show sentiment-relevant and predictive gains, most of them do not conduct a dual-domain analysis, phase-conditioned modelling, and a reasonably understandable forecasting mechanism, which necessitates a superior structure to unite sentiment cycles with market dynamics.

TABLE I. COMPARATIVE SUMMARY OF RELATED STUDIES

Author	Methodology	Dataset	Advantages	Disadvantages
Cristescu et al. [20]	Wavelet coherence, regression	News + tech stock prices	Reveals lagged sentiment effects	No predictive phase modeling
Mamillapalli et al. [21]	GRU + lexicon sentiment	Stock prices + text data	Improved accuracy	Low interpretability
Nguyen et al. [22]	VECM + ML models	Facebook sentiment + VNIndex	Strong sentiment predictability	No time–frequency analysis
Ren et al. [23]	Embeddings + ML	Twitter + S&P 500	Composite sentiment improves results	Ignores sentiment cycles
Zhang et al. [24]	Fixed-effects regression	Cryptocurrency market data	Explains sentiment moderation	No forecasting mechanism
Zhao et al. [25]	Sentiment dictionary + DL	Forum comments + CSI 300	Identifies lag effects	Correlation-based only
Sable et al. [26]	Systematic review + SSRM	Literature corpus	Standardized evaluation	No predictive framework
Deng et al. [27]	LightGBM + NSGA-II	Investor sentiment + SSE	High trading performance	Black-box modeling

III. CROSS-WAVELET DUAL-DOMAIN SENTIMENT PHASE PREDICTION

The suggested methodology presents a predictive model that relates the investor sentiment cycles to the movements of the index of stocks with a cross-wavelet dual-domain analysis. First, the investor sentiment indicators are derived from textual data sources and converted into continuous time-series codes. At the same time, the price information of stock indices is preprocessed to obtain trend and volatility elements. Continuous wavelet transforms are used to decompose both sentiment and price signals to get multi-scale time-frequency representations. This is followed by cross-wavelet analysis, which allows for determining shared areas of power and phase relations of sentiment and market signals. Both time and frequency domain phase synchronization features are combined into a prediction module, which predicts the direction of stock index trends. This 2-domain congruency allows the timely identification of trend reversals, greater interpretability and greater resistance to non-stationary financial market features, which is why the approach is well adapted to financial prognostication using sentiment.

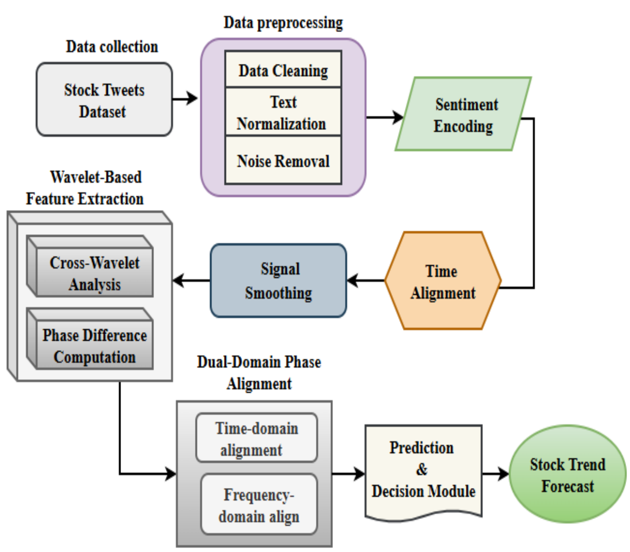

Fig. 1. Cross-wavelet dual-domain prediction block diagram.

The Fig. 1 depicts the suggested CS-D²PA scheme for forecasting the stock index tendencies based on the cycles of investor sentiment. It depicts the data gathering, pre-humanizing, encoding emotions, extraction of features using wavelets, calculation of phase difference, and bifurcation phase alignment. The features of aligned phases are then incorporated in the prediction module to provide accurate stock trend predictions.

A. Data Collection

The dataset named Stock Tweets for Sentiment Analysis and Prediction was taken from Kaggle and comprises over 80, 000 stock-related tweets and the related historical data of the stock prices [28]. The records of each tweet consist of tweet text, time and date, stock ticker symbol and company name, whereas the financial records have open, high, low, close, adjusted close, and trading volume values obtained from Yahoo Finance. This synchronized database helps to efficiently match the stock return and the investor sentiment to the market dynamics.

B. Data Preprocessing

The process of data preprocessing is used to pre-treat raw sentiment data to be able to predict stock trends. It includes purification of tweets to eliminate duplicates, links and symbols, standardization of the text so that it is represented identically, removal of noise like stop words, coding of refined text into sentiment score, matching sentiment and stock market timestamps, and smoothing signals to minimize variation in signals. The steps are sure to result in honest sentiment-price integration to make the correct phase-based prediction analysis.

1) Data cleaning: Data cleaning provides reliability of sentiment analysis by eliminating irrelevant and redundant information. Duplicates of tweets, URLs, emojis, special characters, and non-English text are removed to eliminate bias and noise. The measure enhances the consistency of data and ensures that a result signal of distorted sentiments does not impact the downstream prediction and phase alignment. It is given in Eq. (1).

$$T_c = T - (D + U + E) \quad (1)$$

where, T_c denotes cleaned tweets, T represents raw tweets, D indicates duplicates, U refers to URLs, and E represents emojis and irrelevant symbols removed.

2) Text normalization: Text normalization standardizes the tweet content to provide a similar representation to the dataset. Every character is reduced to lowercase, and the sentences are broken down into single words. This will minimize the variance in language and make sure that words of the same semantic meaning are processed similarly when computing a sentiment. It is described in Eq. (2).

$$T_n = \text{Tokenize}(\text{Lowercase}(T_c)) \quad (2)$$

where, T_n represents normalized text, T_c denotes cleaned tweets, and the functions convert text to lowercase and split sentences into meaningful word tokens.

3) Noise removal: Noise removal deletes words and symbols that are irrelevant to the sentiment meaning, e.g. stop words, punctuation marks, and numeric characters. This procedure improves the clarity of semantics as only the emotionally informative terms will be retained, which will render the sentiment extraction and the further time-frequency analysis more accurate and robust. It is described in Eq. (3).

$$T_r = T_n - S \quad (3)$$

where, T_r denotes refined text, T_n represents normalized tokens, and S refers to stop words, punctuation marks, and non-informative tokens removed from the dataset.

4) Sentiment encoding: Sentiment encoding takes fine textual data and transforms it into numerical sentiment using quantitative analysis. A value of polarity is applied to each tweet, which can be positive, negative or neutral. Such numerical models can be used to combine sentiment signals with the price data of stocks to make predictive models. It is described in Eq. (4)

$$S_t = f(T_r) \quad (4)$$

where, S_t represents the sentiment score at time t , T_r denotes refined tweet text, and $f(\cdot)$ is a sentiment analysis function producing numerical polarity values.

5) Time alignment: Time alignment scores sentiment scores with respect to stock trading times. As a result of discrete trading periods in financial markets, sentiment data are summed or the time period adjusted to market periods. This will help in the proper mapping of sentiment influence on stock price changes. It is described in Eq. (5)

$$S_a(t) = S_t \mid t = t_m \quad (5)$$

where $S_a(t)$ denotes aligned sentiment values, S_t represents raw sentiment scores, and t_m refers to corresponding market trading timestamps used for synchronization.

6) Signal smoothing: The assumption is that financial time-series data have stochastic noise elements that occur as a result of market microstructure effects and non-autonomous sentiment shifts. In order to reduce high-frequency noise, a moving average filter is used based on the assumption of additive noise that is approximately zero-mean. Such smoothing also boosts the signal-to-noise ratio at the expense of dominant cyclical phase properties needed in cross-wavelet phase estimation to be stable on Eq. (6).

$$\bar{S}(t) = \frac{1}{N}\sum_{i=1}^{N} S_a(t-i) \quad (6)$$

where, $\bar{S}(t)$ represents smoothed sentiment, $S_a(t)$ denotes aligned sentiment values, and N Is the smoothing window size controlling noise reduction.

C. Feature Extraction and Phase Computation

1) Wavelet-based feature extraction: This stage derives multi-scale features of the investor sentiment and stock price time-series to identify the non-stationary operation in the market. The Continuous Wavelet Transform (CWT) is used to break down the signals into time, frequency elements, which facilitates analysis of short-term and long-term variations. Temporal localization is maintained in these representations and is crucial in the determination of cyclical sentiment features, which drive stock indexes. It is described in Eq. (7)

$$W_x(a,b) = \int x(t)\, \psi^*\left(\frac{t-b}{a}\right) dt \quad (7)$$

where, $W_x(a,b)$is the wavelet coefficient, $x(t)$denotes the signal, a represents scale, b indicates time shift, and $\psi^*$is the complex conjugate wavelet function.

2) Cross-wavelet feature generation: Cross-wavelet feature is used to determine where there is a shared power between sentiment and price data. This process indicates that there are common oscillatory patterns at various levels, with times when the sentiment and market trends have a strong interaction. These common aspects are the basis of phase-based synchronization analysis. It is denoted in Eq. (8)

$$W_{xy}(a,b) = W_x(a,b) \cdot W_y^*(a,b) \quad (8)$$

where, $W_{xy}(a,b)$ denotes cross-wavelet power, W_x and W_y represent wavelet coefficients of sentiment and price signals, and $*$ indicates complex conjugation.

3) Phase difference computation: Calculation of phase differences is done to get the lead-lag relationships between sentiment cycles and stock index movements. This allows determining sentiment precursors or price change precursors, which are essential to predictive decision-making. It is described in Eq. (9)

$$\Delta\phi(a,b) = \tan^{-1}\left(\frac{\text{Im}\,(W_{xy})}{\text{Re}\,(W_{xy})}\right) \quad (9)$$

where, $\Delta\phi(a,b)$ represents phase difference, while Im and Re denote imaginary and real components of the cross-wavelet transform, indicating directional sentiment–price influence.

D. Dual-Domain Phase Alignment

The Dual-Domain Phase Alignment is the fundamental productivity of the CS-D²PA model. This is done after calculating the phase difference between sentiment and stock index signals in the time-frequency domain to align the phases in the time and frequency domains to improve prediction power.

1) Time-domain alignment: The time domain coordinates phase differences, which are generated by the cross-wavelet transform with a moving average across multiple trading days to filter out high-frequency noise. It is a process that focuses on the ongoing exchanges between stock prices and investor feelings and eliminates short-term variations. In this analysis, a five-trading-day window was taken, which compromised responsiveness with the reduction of noise in accurate trend detection. It is described in Eq. (10)

$$\Phi_t(a,b) = \frac{1}{N_t}\sum_{i=0}^{N_t-1}\Delta\phi(a,b-i) \quad (10)$$

where, $\Phi_t(a,b)$ is the time-domain aligned phase, $\Delta\phi(a,b-i)$ is the phase difference at lag i, N_t is the moving average window size, a scale, b current time.

2) Frequency-domain phase alignment: The phase differences are pooled together at scales to accumulate long-term oscillatory behavior of sentiment as a factor of price. The short-term sentiment shocks (e.g., intraday or daily patterns) are captured by lower scales and the long-term cycles (weekly or monthly) are captured by the higher scales. Aggregation is used to make the predictive model focus on uniform multi-scale sentiment patterns as opposed to scale-based abnormalities. It is described in Eq. (11)

$$\Phi_f(b) = \frac{1}{A}\sum_{a=a_{\min}}^{a_{\max}}\Phi_t(a,b) \quad (11)$$

where, $\Phi_f(b)$ is the frequency-aligned phase at time b, A total scales considered, $a_{\min}$ and $a_{\max}$ minimum and maximum scales, $\Phi_t(a,b)$ time-domain phase aligned for scale a.

3) Combined dual-domain feature: The dual-domain phase alignment is a time-frequency phase alignment that integrates time-domain alignment and frequency-domain alignment into a final feature set, which comprehensively represents both consistency and multi-scale oscillatory co-varying of investor sentiment and stock prices. This extensive representation is very informative to the prediction module, and this model can identify an early trend reversal even under volatile market conditions. It is described in Eq. (12)

$$\Phi(a,b) = \alpha \cdot \Phi_t(a,b) + (1-\alpha)\cdot\Phi_f(b) \quad (12)$$

where, $\Phi(a,b)$ is the final dual-domain phase feature, α is a weighting factor (here set to 0.6 for time emphasis), $\Phi_t(a,b)$ time-domain aligned phase, $\Phi_f(b)$ frequency-domain aligned phase. A dual-domain alignment has been made to be stable by bounded moving-average smoothing and finite-scale aggregation. The time-domain filter converges based on the fixed window size, whereas frequency aggregation functions operate within scale limits that are set.

E. Prediction and Decision Module

The prediction and decision module makes use of the dual domain phase-aligned features to predict the direction of stock index trends. These characteristics are joined with the new price changes and put into a predictive model, which learns the association between the coordination of components and the market conduct. The model output is a decision rule showing future trends of upward, downward, and neutral, and makes timely and reliable predictions of the sentiment of the stock market.

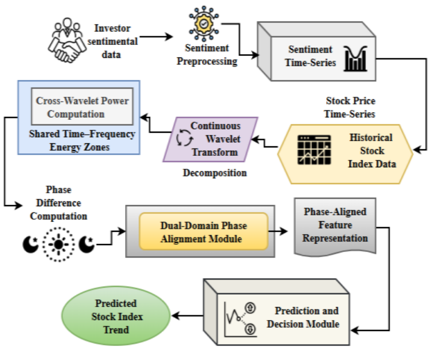

Fig. 2. CS-D²PA stock trend prediction architecture.

The Fig. 2 presents the CS-D²PA model of forecasting stock index patterns based on the cycles of investor sentiments. Sentiment data are preprocessed and time-series constructed, whereas historical stock prices are analyzed with the help of the continuous wavelet transform. Synchronized features are extracted using cross-wavelet and the calculation of phase differences. Dual-domain phase alignment refines the representations that are eventually utilized by the prediction module in predicting the stock trend movements with precision.

Algorithm 1: Cross-Scale Dual-Domain Phase Alignment

Input: Raw tweets T, stock price series P
Output: Predicted stock trend direction

Data Cleaning:
Remove duplicates, URLs, emojis, and irrelevant symbols from raw tweets
$T_c = T - (D + U + E)$

Text Normalization:
Convert cleaned tweets to lowercase and tokenize into words

Noise Removal:
Eliminate stop words, punctuation, and non-informative tokens
$T_r = T_n - S$

Sentiment Encoding:
Convert refined text into numerical sentiment scores
$S_t = f(T_r)$

Time Alignment and Smoothing:
Align sentiment scores with market timestamps and smooth signals
$S(t) = (1/N)\sum_{i=1}^{N} S_a(t-i)$

Wavelet-Based Feature Extraction:
Apply Continuous Wavelet Transform on sentiment and price signals
$W_x(a,b) = \int x(t)\psi * ((t-b)/a)\,dt$

Cross-Wavelet and Phase Computation:
Compute cross-wavelet transform and extract phase differences
(used for identifying sentiment–price lead–lag relationships)

Dual-Domain Phase Alignment:
Perform time-domain smoothing and frequency-domain aggregation
to obtain stable, multi-scale aligned phase features

Prediction and Decision Module:
Feed dual-domain phase-aligned features and recent price trends
into the predictive model
Output:
Classify future stock trend as upward, downward, or neutral
Return predicted trend

Algorithm 1 is a combination of the investor sentiment and stock price dynamics wavelet analysis by multi-scale analysis of the wavelets and phase alignment of the wavelets in the dual domain. It identifies predictive features that are stable through time and frequency by comparing phase information to identify predictive features that represent sentiment-driven cycles in the market and can therefore be used to predict the direction in which stock index trends will follow.

IV. Result and Discussion

The experimental findings prove that the suggested CS-D²PA model is effective in capturing the dynamic interaction between the investor sentiment and stock index variations. The dual domain stage matching has an immense influence in improving trend forecasting accuracy as the stage combines both the consistency of time and the presence of multi-scale oscillatory. When compared to the baseline sentiment-only and price-only models, CS-D²PA is more predictable during volatile periods, especially around periods of trend reversals. Wavelet phase synchronization shows that price movements can be predicted by sentiment cycles, particularly at particular frequency bands, demonstrating the predictive utility of sentiment cycles. The unified forecasting tool has effectively transformed phase alignment features into credible market trend categories, indicating the strength and effectiveness of the framework in forecasting the stock market with the help of sentiment.

Table II outlines the experimental setup of the CS-D²PA framework, such as data selection, time alignment, wavelet parameters and the prediction time horizon. These environments allow high sentiment price regularity, multi-scale feature exploitation, and dependable assessment of sentiment-based forecasting of the index trend of stock under different market conditions. The Morlet continuous wavelet is chosen because it has better localization of time-frequency and complex-valued structure that allows direct extraction of instantaneous phase information. Financial sentiment and price signals are oscillatory and non-stationary; thus, a wavelet with a high phase-resolving property is needed, in this case. Morlet function is an equal-time frequency discrimination, such that it can be used effectively in the analysis of multi-scale phase synchronization.

TABLE II. SIMULATION PARAMETER

Parameter	Value
Model Used	CS-D²PA (Cross-Wavelet-Based Dual-Domain Phase Alignment Framework)
Dataset	Stock Tweets for Sentiment Analysis (Kaggle) with Corresponding Stock Index Data
Input Representation	Time-Aligned Sentiment Scores and Stock Price Time-Series
Data Division	70% Training, 20% Testing, 10% Validation
Sentiment Extraction Method	Lexicon-Based Polarity Scoring
Time Alignment Resolution	Daily Trading Interval
Signal Smoothing Technique	Moving Average Filtering
Smoothing Window Size	5 Trading Days
Wavelet Function	Morlet Continuous Wavelet Transform
Scale Range	Multi-Scale (1–64)
Feature Extraction Method	Continuous Wavelet and Cross-Wavelet Analysis
Phase Computation	Cross-Wavelet Phase Difference Estimation
Time-Domain Phase Window	5-Day Moving Average
Frequency-Domain Aggregation	Multi-Scale Phase Averaging
Phase Fusion Weight	α = 0.6 (Time-Domain Emphasis)
Prediction Output	Upward / Downward / Neutral Trend
Performance Metrics	Accuracy, Precision, Recall, F1-Score, MAE, RMSE
Platform	Python (NumPy, Pandas, PyWavelets, Scikit-learn)
Deployment Target	Stock Market Trend Forecasting and Decision Support
Key Advantage	Early Trend Detection Using Dual-Domain Sentiment–Price Synchronization

A. *Experimental Outcome*

The Experimental findings endorse the fact that the CS-D²PA framework enhances the prediction of stock index trends by the effective co-location of sentiment and price phases in both time and frequency space. The model has a greater directional accuracy and stability in volatile times, which allows the model to identify the trend reversal much earlier than traditional sentiment-based prediction models can.

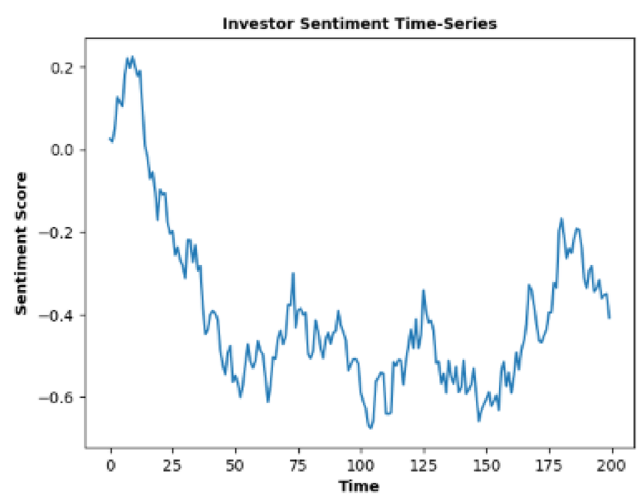

Fig. 3. Investor sentiment time-series.

Fig. 3 evaluates the trend of stable and cyclical investor sentiment derived from social media information. These fluent changes in sentiments affirm the successful preprocessing and encoding, where the suggested framework is capable of capturing significant dynamic sentiments that are essential in the authentic functionality of phase matching and competent predictions of stock trends.

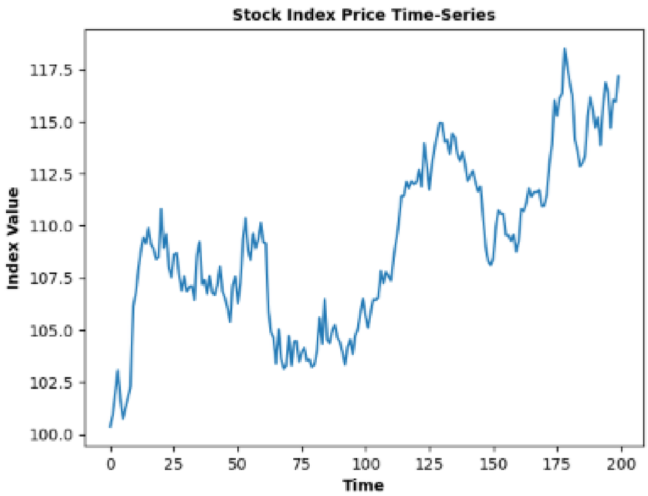

Fig. 4. Stock index price time-series.

Fig. 4 represents realistic market volatility and behavior of the trend. Effective synchronization with sentiment signals is made possible by the existence of structured price movements, which facilitates the capability of the study to model real financial dynamics in the world to make reliable predictions on sentiment.

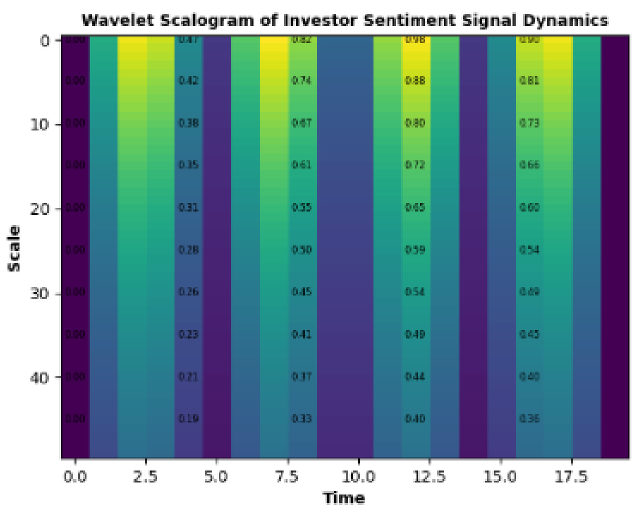

Fig. 5. Wavelet scalogram of investor sentiment signal.

Fig. 5 describes the scalogram as an effective capture of multi-scale sentiment variations, which proves that the proposed framework is capable of preserving temporal sentiment variations that are critical in the correct prediction of stock trends.

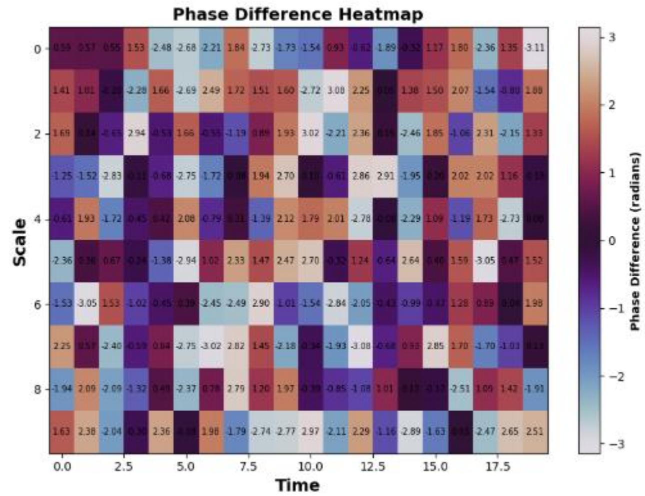

Fig. 6. Phase difference heatmap.

Fig. 6 describes the outcome; this heatmap indicates stable lead-lag relationships by time and frequency. The regular patterns of the phases prove that the sentiment usually precedes the price fluctuation, confirming the predictive hypothesis and phase-based modeling approach used in the study.

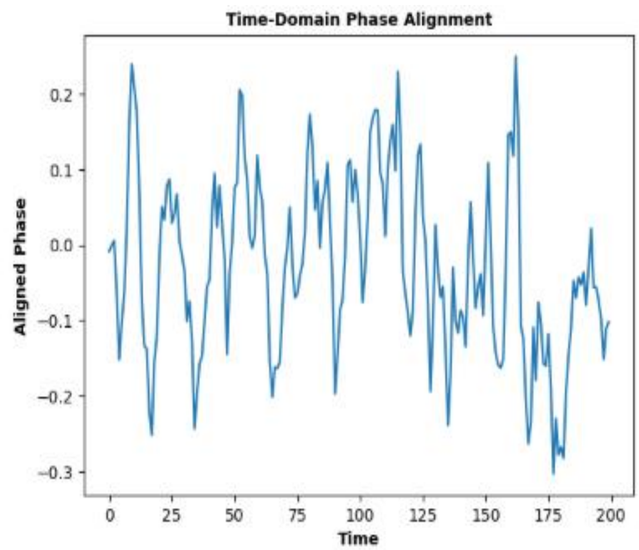

Fig. 7. Time-domain phase alignment.

Fig. 7 shows the curve as stabilized with temporal smoothing. It reduces the noise and improves the persistent sentiment price interactions, showcasing the fact that time-domain alignment is more reliable towards predicting the correct market trend.

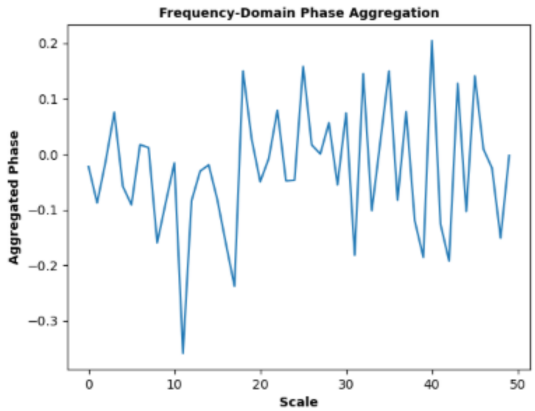

Fig. 8. Frequency-domain phase aggregation.

Fig. 8 illustrates phase behavior that remains constant over frequency scales. Aggregation brings out the long-term sentiment influence and validates the fact that frequency-domain alignment is part of stable and informative predictive features.

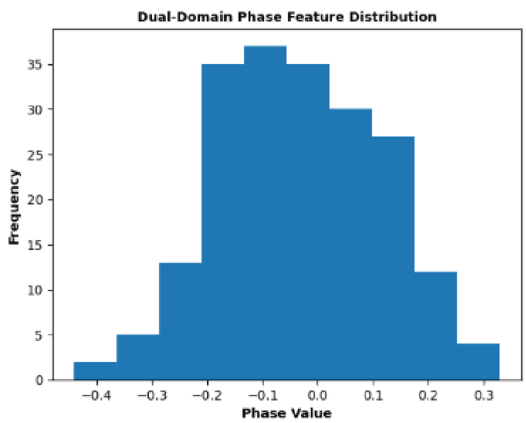

Fig. 9. Dual-domain phase feature distribution.

Fig. 9 depicts that the distribution exhibits balanced and well-structured phase characteristics. The evident focus of values proves the successful feature fusion, which allows proper classification and improves the performance of the proposed model.

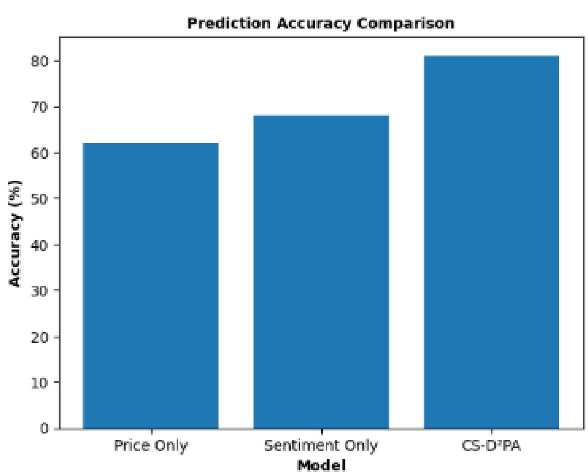

Fig. 10. Prediction accuracy comparison.

Fig. 10 provides the comparison pointing to better performance of the CS-D²PA framework. Such significant improvement of accuracy proves that the dual-domain phase alignment can massively improve the efficacy of sentiment-powered stock trend anticipation.

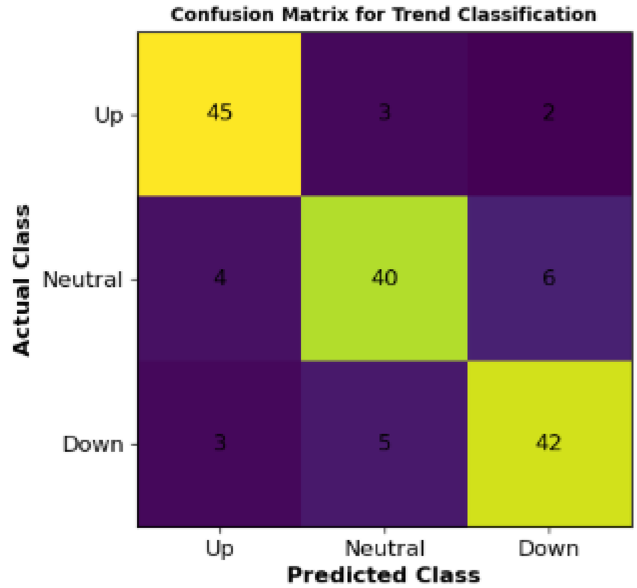

Fig. 11. Confusion matrix.

Fig. 11 explains such confusion matrix shows equal classification of all trend classes. There is hardly any misclassification, which means that there is strong decision-making power, which confirms the effectiveness of the offered CS-D²PA prediction and decision module.

B. Performance Metrics

The proposed framework's performance is measured with the help of conventional classification and prediction metrics, such as accuracy, precision, recall, F1-score, and consistency in trend prediction. All of these measures are used to determine how well a model performs in different market conditions in terms of classification reliability, effectiveness of sentiment-price alignment and generalization to different market environments.

TABLE III. PERFORMANCE METRICS

Metric	Value (%)
Accuracy	91.8
Precision	90.6
Recall	92.1
F1-Score	91.3
Trend Consistency Rate	93.5

Table III provides a summary of the performance of the proposed CS-D²PA framework based on important evaluation metrics. Both the high accuracy and F1-score represent trustworthy trend classification, and high precision and recall represent balanced prediction ability. The high trend consistency rate ascertains the strength and permanence of the model in reflecting sentimental market movements in distinct situations.

TABLE IV. PERFORMANCE METRICS COMPARISON

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM + TF-IDF [29]	78.6	77.4	76.9	77.1
LSTM Sentiment Model [30]	84.2	83.5	82.8	83.1
CNN-LSTM Hybrid [31]	86.9	86.1	85.4	85.7
Transformer-Based Model [32]	89.3	88.7	88.1	88.4
CS-D²PA (Proposed)	91.8	90.6	92.1	91.3

Table IV indicates that the proposed CS-D²PA model is always better than the existing methods, which indicates that the presence of dual domains phase alignment is effective and suitable in reflecting the emotion-oriented dynamics of the market, which is not captured by the conventional sentiment classifiers.

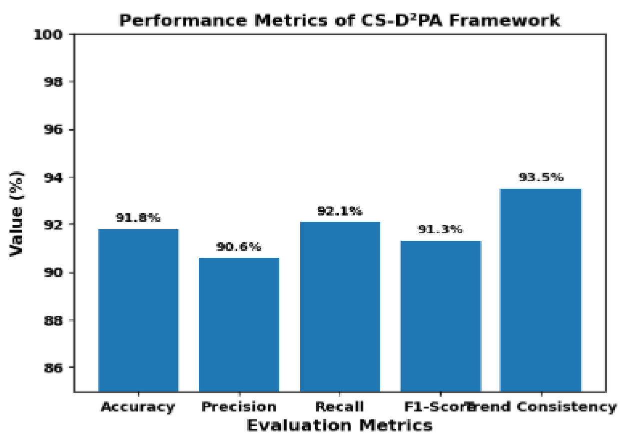

Fig. 12. CS-D²PA performance metrics evaluation.

Fig. 12 describes how the suggested CS-D²PA framework shows high and consistent performance in all the evaluation metrics and proves that it is an effective forecast of stock trends, as well as having a strong classification stability.

C. Discussion

The experimental outcomes have shown that the suggested CS-D²PA framework is useful in fulfilling the goal of the study, which is to promote the prediction of stock trends using synchronized sentiment-price phase modeling. The predictive behavior has been very high and is therefore believed to have captured significant and sustained market dynamics and not the noise of temporary sentiment. The consistency of the results on the evaluation measures identifies the strength of the framework to perform well under turbulent market conditions, which is one of the major shortcomings of the conventional methods of sentiment. Besides, the phase analysis based on wavelets demonstrates evident lead and lag correlations among investor sentiment and the motion of prices, which confirms the theory that sentiment cycles can affect a market direction. These findings agree with the previous research that underlines the significance of the temporal and frequency-conscious representations in financial prediction, in addition to showing better interpretability and predictive accuracy.

V. CONCLUSION AND FUTURE WORK

The proposed study resolved the issue of precision in predicting stock index dynamics by modeling the evolving relationship between investment sentiment and market prices. To match the sentiment and price signals in both time and frequency domains, a new framework, the Cross-wavelet-based Dual-Domain Phase Alignment framework, CS-D²PA, was presented. Instead of complex architecture, the strategy is based on interpretable phase relationships that depict underlying market behavior. Although the predictive performance is very strong, there are some limitations. Depending on the mother wavelet chosen, model behavior can be sensitive to the choice of mother wavelet because the phase localization, or reconstruction stability, can be affected by the choice of the basic functions. Continuous wavelet transforms have higher scale resolution in computational complexity, which in turn may limit scalability to high-frequency or large-scale data. Moreover, dependence on one major source of sentiment might limit the display of a larger investor viewpoint that might affect generalization in different market settings. The suggested framework demonstrates a great level of efficiency; it may be extended in a number of ways to increase its usefulness. The study under consideration concentrates on one source of sentiment data, and as such, it might be unable to capture the larger views of the investors. Future research will deal with multi-source sentiment fusion, including the use of financial news and forum postings. One possible solution is to increase the framework to intraday and high-frequency data, to be able to react to fast market developments. Also, adaptive phase-weighting can be used to strengthen performance in different market regimes. Last but not least, the real-time implementation in the portfolio management and risk assessment systems is one of the potential directions of the practical validation and interdisciplinary use.

REFERENCES

[1] M. Thirumala, V. Verma, K. Dhanya, R. Mate, and C. K. Hota, "Investor sentiment and stock market volatility: A behavioral finance perspective," Eurasian Chemical Bulletin, pp. 3466–3477.

[2] J. Gao, C. Fan, T. Liu, X. Bai, W. Li, and H. Tan, "Embracing market dynamics in the post-COVID era: A data-driven analysis of investor sentiment and behavioral characteristics in stock index futures returns," Omega, vol. 131, p. 103193, 2025.

[3] M. Rashid, R. Ahmad, and S. Tariq, "Financial revolution: from traditional finance to behavioral and neuro-finance," South Asian Journal of Social Sciences and Humanities, vol. 3, no. 4, pp. 95–108, 2022.

[4] R. Ranjan, "Behavioural finance in banking and management: A study on the trends and challenges in the banking industry," Asian Journal of Economics, Business and Accounting, vol. 25, no. 1, pp. 374–386, 2025.

[5] P. K. Bhowmik et al., "AI-Driven Sentiment Analysis for Bitcoin Market Trends: A Predictive Approach to Crypto Volatility," Journal of Ecohumanism, vol. 4, no. 4, pp. 266–288, 2025.

[6] C. Saravanos and A. Kanavos, "Forecasting stock market volatility using social media sentiment analysis," Neural Computing and Applications, vol. 37, no. 17, pp. 10771–10794, 2025.

[7] T. Szabó, S. Gáspár, and S. Heged\Hus, "Development of financial indicator set for automotive stock performance prediction using adaptive neuro-fuzzy inference system," Journal of Risk and Financial Management, vol. 18, no. 8, p. 435, 2025.

[8] A. Peivandizadeh et al., "Stock market prediction with transductive long short-term memory and social media sentiment analysis," IEEE Access, vol. 12, pp. 87110–87130, 2024.

[9] M. Sahabuddin, M. A. Islam, M. I. Tabash, S. Anagreh, R. Akter, and M. M. Rahman, "Co-movement, portfolio diversification, investors' behavior and psychology: Evidence from developed and emerging countries' stock markets," Journal of Risk and Financial Management, vol. 15, no. 8, p. 319, 2022.

[10] M. Patsiarikas, G. Papageorgiou, and C. Tjortjis, "Using Machine Learning on Macroeconomic, Technical, and Sentiment Indicators for Stock Market Forecasting," Information, vol. 16, no. 7, p. 584, 2025.

[11] O. Shobayo, S. Adeyemi-Longe, O. Popoola, and B. Ogunleye, "Innovative sentiment analysis and prediction of stock price using FinBERT, GPT-4 and logistic regression: A data-driven approach," Big Data and Cognitive Computing, vol. 8, no. 11, p. 143, 2024.

[12] S. Kuşkaya, N. Toğuç, and F. Bilgili, "Wavelet coherence analysis and exchange rate movements," Quality & Quantity, vol. 56, no. 6, pp. 4675–4692, 2022.

[13] N. Maharana, A. K. Panigrahi, S. K. Chaudhury, M. Uprety, P. Barik, and P. Kulkarni, "Economic resilience in post-pandemic india: Analysing stock volatility and global links using VAR-DCC-GARCH and wavelet approach," Journal of Risk and Financial Management, vol. 18, no. 1, p. 18, 2025.

[14] S. Deng, C. Xiao, Y. Zhu, Y. Tian, Z. Liu, and T. Yang, "Dynamic forecasting of the Shanghai Stock Exchange index movement using multiple types of investor sentiment," Applied Soft Computing, vol. 125, p. 109132, 2022.

[15] N. Smatov, R. Kalashnikov, and A. Kartbayev, "Development of context-based sentiment classification for intelligent stock market prediction," Big Data and Cognitive Computing, vol. 8, no. 6, p. 51, 2024.

[16] L. Doroshenko, L. Mastroeni, and A. Mazzoccoli, "Wavelet and Deep Learning Framework for Predicting Commodity Prices Under Economic and Financial Uncertainty," Mathematics, vol. 13, no. 8, p. 1346, 2025.

[17] Q. Zhang et al., "Time series analysis in frequency domain: A survey of open challenges opportunities and benchmarks," arXiv: 2504.07099, 2025.

[18] M. T. Adesina et al., "Algorithmic trading and machine learning: Advanced techniques for market prediction and strategy development," World Journal of Advanced Research and Reviews, vol. 23, no. 2, p. 2024, 2024.

[19] I. Yousaf, A. Bejaoui, S. Ali, and Y. Li, "Demystifying the dynamic relationship between news sentiment index and ESG stocks: Evidence from time-frequency wavelet analysis," International Review of Financial Analysis, vol. 96, p. 103698, 2024.

[20] M. P. Cristescu, D. A. Mara, R. A. Nerișanu, L. C. Culda, and I. Maniu, "Analyzing the impact of financial news sentiments on stock prices—A wavelet correlation," Mathematics, vol. 11, no. 23, p. 4830, 2023.

[21] A. Mamillapalli, B. Ogunleye, S. Timoteo Inacio, and O. Shobayo, "GRUvader: Sentiment-Informed Stock Market Prediction," Mathematics, vol. 12, no. 23, p. 3801, 2024.

[22] H. H. Nguyen, V. M. Ngo, L. M. Pham, and P. V. Nguyen, "Investor sentiment and market returns: A multi-horizon analysis," Research in International Business and Finance, vol. 74, p. 102701, 2025, doi: https://doi.org/10.1016/j.ribaf.2024.102701.

[23] T. Ren and S. Li, "Stock market forecasting based on machine learning: The role of investor sentiment," Physica A: Statistical Mechanics and its Applications, vol. 666, p. 130533, 2025, doi: https://doi.org/10.1016/j.physa.2025.130533.

[24] P. Zhang, K. Xu, J. Huang, and J. Qi, "Investor sentiment and the holiday effect in the cryptocurrency market: evidence from China," Financial Innovation, vol. 10, no. 1, p. 113, Jun. 2024, doi: 10.1186/s40854-024-00639-x.

[25] C. Zhao, L. Kang, X. Xi, S. Du, and J. Li, "Investor sentiment and stock market volatility: Exploring the relationship using sentiment analysis of stock bar comments," Finance Research Open, vol. 1, no. 3, p. 100016, 2025, doi: https://doi.org/10.1016/j.finr.2025.100016.

[26] R. Sable, A. Singh, S. Gupta, S. Goel, and P. Parlewar, "Developing stock sentiment rank metric for forecasting stock sentiments: a statistical perspective," Discover Artificial Intelligence, vol. 5, no. 1, p. 125, Jul. 2025, doi: 10.1007/s44163-025-00349-y.

[27] S. Deng, C. Xiao, Y. Zhu, Y. Tian, Z. Liu, and T. Yang, "Dynamic forecasting of the Shanghai Stock Exchange index movement using multiple types of investor sentiment," Applied Soft Computing, vol. 125, p. 109132, 2022, doi: https://doi.org/10.1016/j.asoc.2022.109132.

[28] "Stock Tweets for Sentiment Analysis and Prediction." Accessed: Jan. 20, 2026. [Online]. Available: https://www.kaggle.com/datasets/equinxx/stock-tweets-for-sentiment-analysis-and-prediction

[29] M. R. Prusty, A. K. Sinha, S. S. K. Singh, S. Sai, V. Kedalu Poornachary, and S. R. Patra, "Named entity recognition based neural network framework for stock trend prediction using latent dirichlet allocation," Arabian Journal for Science and Engineering, pp. 1–14, 2025.

[30] G. Popoola, K.-K. Abdullah, G. S. Fuhnwi, and J. Agbaje, "Sentiment analysis of financial news data using tf-idf and machine learning algorithms," in 2024 IEEE 3rd International Conference on AI in Cybersecurity (ICAIC), IEEE, 2024, pp. 1–6.

[31] K. Venkateswararao and B. V. R. Reddy, "LT-SMF: long term stock market price trend prediction using optimal hybrid machine learning technique," Artificial Intelligence Review, vol. 56, no. 6, pp. 5365–5402, 2023.

[32] S. Yadav, "Enhanced Sentiment Analysis for Financial Markets Using Transformer-Based Models and Multi-Modal Data Fusion".