IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 9 August 2025, accepted 30 August 2025, date of publication 2 September 2025, date of current version 11 September 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3605522

RESEARCH ARTICLE

HSIF: A Transformer-Based Cross-Attention Framework for Cryptocurrency Trend Forecasting via Multimodal Sentiment–Market Fusion

SAEED MOHAMMADI DASHTAKI[1], MEHDI HOSSEINI CHAGAHI[1], AEIN BAHADORI[2], BEHZAD MOSHIRI[1,3], (Senior Member, IEEE), MD. JALIL PIRAN[4], (Senior Member, IEEE), AND ALLAHYAR MONTAZERI[5], (Member, IEEE)

[1]School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran
[2]Department of Information Engineering, University of Padova, 35122 Padua, Italy
[3]Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[4]Department of Computer Science and Engineering, Sejong University, Seoul 05006, South Korea
[5]School of Engineering, Lancaster University, LA1 4YW Lancaster, U.K.

Corresponding authors: Allahyar Montazeri (a.montazeri@lancaster.ac.uk) and Behzad Moshiri (moshiri@ut.ac.ir)

ABSTRACT Cryptocurrency markets are highly volatile and sentiment-driven, posing challenges to traditional forecasting methods. This paper presents Hard and Soft Information Fusion (HSIF), a novel Transformer-based dual-stream model that combines market data and social sentiment using Financial Bidirectional Encoder Representations from Transformers (FinBERT), a financial sentiment analysis tool, and a bidirectional cross-attention mechanism. Evaluations on multi-year Bitcoin data show that HSIF achieves 97.48% accuracy and a 26.64% return, outperforming Long Short-Term Memory (LSTM)-based and other multimodal models. The results highlight the effectiveness of domain-specific sentiment embeddings and cross-modal attention in enhancing trend prediction accuracy for volatile cryptocurrency markets.

INDEX TERMS Cryptocurrency price forecasting, transformer-based models, multimodal information fusion, cross-attention mechanism, natural language processing, sentiment analysis.

I. INTRODUCTION

Cryptocurrencies have emerged as a transformative force in the global financial landscape, attracting substantial attention from investors and researchers. By early 2022, the market capitalization of cryptocurrencies reached nearly two trillion dollars, signifying the growing adoption and integration of digital assets into mainstream financial systems [1], [2]. Among various cryptocurrencies, Bitcoin, introduced in 2008 as the first decentralized digital currency, continues to hold a dominant position, both as a store of value and a peer-to-peer payment system [3], [4], [5].

The decentralized nature of Bitcoin, combined with the volatility inherent in cryptocurrency markets, presents both opportunities and risks for investors. In early 2022, Bitcoin represented approximately 40% of the total cryptocurrency market capitalization, underscoring its significant market influence. However, the challenge of accurately predicting price movements in such a volatile market remains critical for making informed investment decisions [1], [3]. Recent studies have also highlighted the need for improved forecasting techniques that consider multiple factors influencing the markets, such as macroeconomic variables and sentiment analysis [6], [7], [8].

Traditional methods for cryptocurrency price prediction typically rely on quantitative data, such as historical prices and technical indicators [9], [10], [11], [12]. Recent research, however, has emphasized the importance of incorporating qualitative data, such as investor sentiment derived from social media, news articles, and other external sources, in improving predictive models. This shift towards integrating both quantitative and qualitative data is essential, as price

The associate editor coordinating the review of this manuscript and approving it for publication was Arianna D'Ulizia.

movements in cryptocurrency markets are often influenced by a combination of objective factors and market psychology [13], [14], [15], [16], [17].

Recent advancements in artificial intelligence (AI) have significantly improved the accuracy of predictive models by enabling them to capture complex patterns within large and diverse datasets. AI has been successfully applied to various financial tasks, including algorithmic trading, risk management, fraud detection, and credit scoring. The application of AI to cryptocurrency price prediction has similarly gained attention, with promising results achieved using both machine learning (ML) algorithms and more sophisticated deep learning (DL) approaches [18], [19], [20], [21], [22], [23], [24].

A promising development in the field of AI is the use of Transformer-based architectures, which excel at capturing long-range dependencies in sequential data. These architectures, initially developed for Natural Language Processing (NLP) tasks, have shown great potential in various domains, including financial market prediction. The incorporation of NLP techniques, such as sentiment analysis, into cryptocurrency price forecasting models offers the ability to capture market sentiment from social media platforms and news outlets, further enhancing prediction accuracy [25], [26], [27], [28].

A clear distinction exists between data fusion and information fusion. While data fusion typically involves the integration of raw or low-level features from different sources [29], information fusion refers to the combination of processed, semantically rich representations [30]. The proposed Hard and Soft Information Fusion (HSIF) framework adopts an information fusion strategy by merging high-level features from both market indicators and sentiment embeddings through bidirectional cross-attention. This enables the model to learn deeper dependencies across modalities, which is essential for robust prediction in sentiment-driven markets.

Cryptocurrency markets are characterized by extreme volatility, influenced by both quantitative market data and qualitative factors such as news. However, current forecasting models primarily rely on technical indicators or simplistic combinations of sentiment analysis and technical indicators, often failing to capture the intricate interactions between these factors. While DL models, such as Long Short-Term Memory (LSTM)-based and Transformer-based architectures, have shown promise, they tend to either overlook sentiment or treat it in a simplistic, one-dimensional manner. To address these challenges, we propose a novel Transformer-based framework, HSIF, which fuses hard (quantitative) and soft (qualitative) data through a mutual attention mechanism. By leveraging NLP with a Bidirectional Encoder Representations from Transformers (BERT)-based model for sentiment analysis and integrating it with technical indicators, our model not only enhances forecasting accuracy but also improves robustness and interpretability. Extensive experiments demonstrate that HSIF significantly outperforms existing models in both classification accuracy and financial returns, providing a more comprehensive approach to cryptocurrency trend prediction.

In summary, the contributions of this study are as follows:

- A novel Transformer-based cross-attention architecture is proposed, capable of fusing heterogeneous data modalities, specifically, structured quantitative market indicators and unstructured textual sentiment, to model complex temporal and semantic dependencies in cryptocurrency trend prediction.
- A financial-domain NLP pipeline is implemented using financial BERT (FinBERT) to extract sentiment features from social media and news content. These sentiment-aware representations complement technical indicators by capturing investor psychology relevant to market trends.
- Extensive experiments validate the proposed model's superiority over traditional and DL baselines, achieving state-of-the-art predictive accuracy and demonstrating real-world profitability in simulated trading, thus highlighting the model's viability for algorithmic investment strategies in highly volatile digital asset markets.

The remainder of this paper is organized as follows. Section II presents the proposed methodology, detailing the dual-stream attention-based model and the fusion of hard and soft information. Section III describes the experimental setup, including datasets, evaluation metrics, and results. Section IV analyzes the findings, highlights the architectural contributions of HSIF in contrast to related multimodal and DL models. Finally, Section V discusses the findings and concludes with directions for future research.

II. PROPOSED HSIF MODEL

Cryptocurrency markets are influenced by both hard data or quantitative signals and soft data or qualitative factors (e.g., investor sentiment from news and social media). To address this, a novel NLP-driven dual-stream architecture, named the HSIF model, is proposed. The model employs Transformer-based encoders to process structured technical indicators and unstructured text-based sentiment in parallel, with bidirectional cross-attention for effective multimodal fusion.

This section details the model design, preprocessing pipeline, and integration of natural language information to enhance the forecasting of the cryptocurrency price movements. The overall architecture is shown in Figure 1, and Table 1 defines the notations used throughout.

A. PROBLEM DEFINITION

The task is formulated as a binary classification problem: predicting the directional movement of the closing price on the next trading day. Let C_t denote the closing price on day t, then:

$$y_{t+1} = \begin{cases} 1 & \text{if } C_{t+1} > C_t \quad \text{(upward)} \\ 0 & \text{otherwise} \quad \text{(downward)}, \end{cases} \tag{1}$$

TABLE 1. Summary of notations used throughout the manuscript.

Notation	Description
C_t	The closing price of the cryptocurrency on day t.
O_t	The opening price of the cryptocurrency on day t.
H_t	The highest price of the cryptocurrency on day t.
L_t	The lowest price of the cryptocurrency on day t.
$Volume_t$	The volume of the cryptocurrency traded on day t.
N	The entire time period.
T	The time window before the movement prediction day.
$\mathbf{h}_t$	The hard information on trading day t.
$\mathbf{s}_t$	The soft information on trading day t.
$\mathbf{H}_t$	The hard information on a T period on trading day t.
$\mathbf{S}_t$	The soft information on a T period on trading day t.
y_t	The price movement label on trading day t.
$\hat{y}_t$	The predicted price movement label on trading day t.

where the label y_{t+1} indicates the movement of the cryptocurrency price, where a value of "1" shows a positive movement (upward), and a value of "0" indicates a negative movement (downward) [10], [15].

B. PREPROCESSING

1) TEXTUAL DATA PREPROCESSING (SOFT INFORMATION)

Incorporating sentiment from social media is pivotal in modeling the psychological behaviors that dominate cryptocurrency markets. To this end, this study leverages soft information extracted from the X platform (formerly Twitter) [31], where domain-specific user accounts continuously disseminate cryptocurrency-related content. Unlike general user-generated noise, only curated, high-frequency accounts were selected to ensure semantic reliability and contextual alignment with price movement patterns.

The raw textual stream is subjected to a rigorous NLP pipeline designed to maximize signal quality and prepare the data for DL-based sentiment fusion:

1) **Noise Removal and Standardization:** All tweets undergo initial de-noising where user mentions are replaced by the token `@user`, and hyperlinks are masked as `http`. This prevents sentiment misclassification caused by entities or irrelevant links and ensures token uniformity.
2) **Domain-Specific Tokenization:** Each tweet is lowercased and tokenized using the FinBERT tokenizer, a BERT-based tokenizer pretrained specifically on financial language corpora.
3) **Transformer-Based Sentiment Encoding:** The tokenized tweets are passed through the FinBERT model [32], which outputs a probability distribution over three sentiment classes: positive, neutral, and negative. FinBERT, unlike general-purpose BERT or even RoBERTa and CryptoBERT, is fine-tuned on financial sentiment data, including analyst reports and earnings calls, allowing for a higher semantic fidelity in financial domains.
4) **Temporal Aggregation of Sentiment Vectors:** Sentiment is aggregated at the daily level by computing the arithmetic mean of the sentiment probabilities across all tweets posted on the same trading day. This results in a time-aligned sentiment vector $\mathbf{s}_t = [s_t^{\text{pos}}, s_t^{\text{neu}}, s_t^{\text{neg}}] \in \mathbb{R}^{1\times 3}$, which encapsulates the prevailing market sentiment for that day. Importantly, this aggregation mechanism acts as a natural denoising filter: any isolated or manipulative content, such as fake news or outlier sentiment, is diluted in the presence of a large volume of reliable tweets from curated sources. As a result, the overall daily sentiment signal becomes inherently robust against misinformation, reducing the need for explicit fake news detection mechanisms.

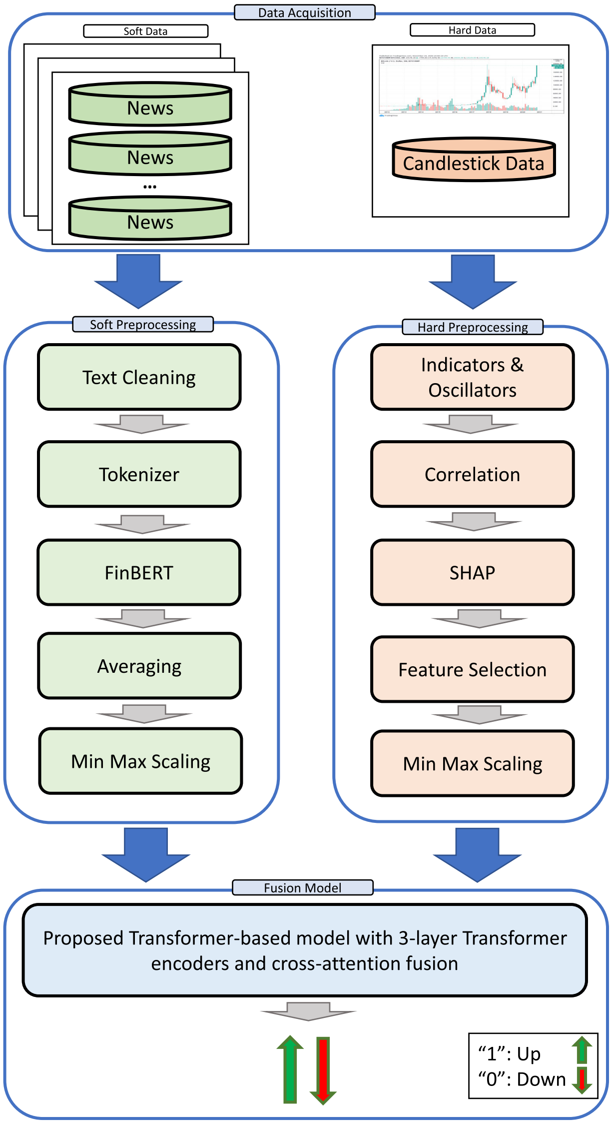

FIGURE 1. The proposed hard and soft information fusion framework for cryptocurrency price movement prediction.

5) **Sequential Sentiment Matrix Formation:** For each prediction step, a rolling time window of length T days is applied to construct the soft information matrix:

$$\mathbf{S}_t = [\mathbf{s}_{t-T+1}, \ldots, \mathbf{s}_t] \in \mathbb{R}^{T \times 3}, \tag{2}$$

This matrix is passed into the soft modality stream of the Transformer encoder to preserve sentiment chronology and capture temporal mood shifts.

Preliminary experiments using other transformer-based sentiment classifiers, such as RoBERTa [33] and CryptoBERT [34], confirmed that FinBERT provides consistently superior performance in both predictive accuracy and financial return. Its domain-tuned architecture proves highly effective in capturing nuanced financial sentiment, particularly in volatile environments like cryptocurrency markets.

The sentiment matrix $\mathbf{S}_t$ becomes a key driver in the fusion process, enabling the model to incorporate qualitative insights alongside technical indicators for more informed and psychologically aware price forecasting.

2) TECHNICAL INDICATOR PROCESSING AND FEATURE SELECTION (HARD INFORMATION)

In modern financial time series modeling, quantitative or ''hard'' data, including historical prices and engineered technical indicators, forms the structural backbone of predictive models. To comprehensively represent the behavior of the cryptocurrency market, this study extracts a diverse set of 52 technical features from historical trading data of the cryptocurrency. These include momentum oscillators, volatility bands, volume-based metrics, and trend indicators, which collectively offer a robust snapshot of market dynamics.

The historical data are sourced from Yahoo Finance [35], a widely trusted accessible platform. For each trading day t, raw market attributes such as opening price O_t, closing price C_t, highest price H_t, lowest price L_t, and traded volume $Volume_t$ are retrieved. These form the basis for computing a comprehensive array of technical indicators. A complete list of the technical indicators and their mathematical definitions is provided in the supplementary material.

a: FEATURE REDUNDANCY REDUCTION VIA CORRELATION FILTERING

Although a wide range of indicators enriches the feature space, they also introduce high inter-feature correlation, which can degrade model performance through redundancy and multicollinearity. To address this, a Pearson correlation analysis is conducted across all 52 initial features. The correlation coefficient r_{xy} between features x and y is computed as:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{3}$$

where $\bar{x}$ and $\bar{y}$ represent the mean values. Features with an absolute correlation coefficient exceeding 0.95 are deemed redundant. This filtering process reduces the dimensionality from 52 to 36 features while retaining representative signal diversity.

b: INTERPRETABLE FEATURE SELECTION

To further refine the feature space and prioritize variables with the highest predictive contribution, SHapley Additive exPlanations (SHAP) values are computed. SHAP provides a model-agnostic, interpretable quantification of each feature's marginal impact on prediction outcomes [36]. The top 30 features, selected based on average absolute SHAP value ranking, are retained for modeling. Figure 2 visualizes the ranked importance of the final feature set.

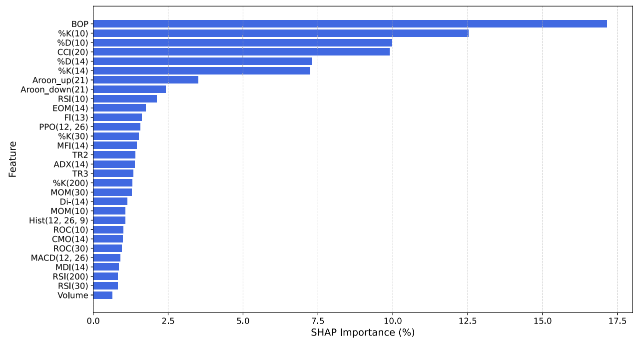

FIGURE 2. Feature importance scores derived from SHAP analysis, highlighting the top 30 most informative hard information.

Table 2 summarizes the selected indicators, their descriptions, and key parameters (e.g., window lengths or smoothing factors). These features capture a multidimensional view of market behavior, including trend strength, volatility shifts, price momentum, and investor pressure, all of which are critical for price direction forecasting.

c: TEMPORAL CONSTRUCTION OF HARD FEATURE MATRICES

Once filtered and selected, these 30 hard features are structured into temporal sequences using a rolling window approach. For each trading day t, the 30-dimensional feature vector is defined as:

$$\mathbf{h}_t = [BOP_t, \ldots, Volume_t] \in \mathbb{R}^{1 \times 30}, \tag{4}$$

and the sequential input over a fixed lookback horizon of T days is constructed as:

$$\mathbf{H}_t = [\mathbf{h}_{t-T+1}, \ldots, \mathbf{h}_t] \in \mathbb{R}^{T \times 30}, \tag{5}$$

where T is the temporal context window, empirically set to 60 in this study. These matrices feed directly into the hard-stream Transformer encoder, enabling temporal attention learning across historical market patterns.

By combining advanced correlation pruning and SHAP-driven interpretability, the hard data pipeline delivers

TABLE 2. Overview of the top 30 selected technical indicators (hard features) used for Bitcoin trend forecasting.

Feature	Description	Parameters
BOP	Balance of Power	–
$\%K(n)$	Stochastic Oscillator %K	10, 14, 30, 200
$\%D(n)$	Stochastic Oscillator %D	10, 14
$CCI(n)$	Commodity Channel Index	20
$Aroon(n)$	Aroon Up/Down	21
$RSI(n)$	Relative Strength Index	10, 30, 200
$EOM(n)$	Ease of Movement	14
$FI(n)$	Force Index	13
$PPO(n_1, n_2)$	Percentage Price Oscillator	(12, 26)
$MFI(n)$	Money Flow Index	14
$TR2$, $TR3$	True Range Components	–
$ADX(n)$	Avg. Directional Movement Index	14
$MOM(n)$	Momentum	10, 30
$DI^-(n)$	Negative Directional Indicator	14
$Hist(n_1, n_2, n_s)$	MACD Histogram	12, 26, 9
$ROC(n)$	Rate of Change	10, 30
$CMO(n)$	Chande Momentum Oscillator	14
$MACD(n_1, n_2)$	MACD Line	(12, 26)
$MDI(n)$	Minus Directional Indicator	14
Volume	Daily Trade Volume	–

a compact, information-rich feature space that enhances learning efficiency while preserving semantic clarity.

C. PROPOSED TRANSFORMER-BASED MULTIMODAL TREND PREDICTION MODEL

This work introduces a novel dual-stream Transformer-based architecture that leverages bidirectional cross-attention for multimodal temporal fusion. The model is designed to effectively learn long-range dependencies within and across quantitative (hard) and qualitative (soft) modalities, enabling enhanced trend prediction in cryptocurrency markets.

1) MULTIMODAL INPUT STRUCTURE

Given a time window of size $T = 60$, the model simultaneously ingests:

- $\mathbf{H}_t \in \mathbb{R}^{T\times 30}$: a matrix of 30 engineered financial indicators across T days, capturing historical price behavior, momentum, volume dynamics, and volatility patterns.
- $\mathbf{S}_t \in \mathbb{R}^{T\times 3}$: a matrix of sentiment vectors derived from FinBERT, encoding average daily scores for positive, neutral, and negative polarity extracted from financial social media discourse.

Each modality is independently normalized using min-max scaling to mitigate scale disparities. The dataset is chronologically partitioned into training (70%), validation (15%), and test (15%) sets, ensuring no temporal leakage.

2) TEMPORAL ENCODING

To preserve sequential context, both inputs undergo positional encoding:

$$\mathbf{H}_t^{(0)} = \mathrm{PosEnc}(\mathbf{H}_t), \tag{6}$$
$$\mathbf{S}_t^{(0)} = \mathrm{PosEnc}(\mathbf{S}_t). \tag{7}$$

3) DUAL-STREAM TRANSFORMER ENCODERS

The encoded matrices are separately passed through three stacked Transformer encoder layers per modality:

$$\mathbf{H}_t^{(l)}\mathrm{TransformerEncoder}^{(l)}(\mathbf{H}_t^{(l-1)}), \quad l = 1, 2, 3, \tag{8}$$
$$\mathbf{S}_t^{(l)} = \mathrm{TransformerEncoder}^{(l)}(\mathbf{S}_t^{(l-1)}), \quad l = 1, 2, 3, \tag{9}$$

yielding final latent representations $\tilde{\mathbf{H}}_t$ and $\tilde{\mathbf{S}}_t$.

4) BIDIRECTIONAL CROSS-ATTENTION FUSION

To capture inter-modal dynamics, two cross-attention layers are introduced:

$$\mathbf{C}_{\mathrm{HS}} = \mathrm{MultiHeadAttention}(\tilde{\mathbf{H}}_t, \tilde{\mathbf{S}}_t), \tag{10}$$
$$\mathbf{C}_{\mathrm{SH}} = \mathrm{MultiHeadAttention}(\tilde{\mathbf{S}}_t, \tilde{\mathbf{H}}_t), \tag{11}$$

where each stream attends to the other to identify relevant interdependencies.

5) MULTIMODAL FEATURE AGGREGATION AND FUSION

The outputs of the cross-attention modules are condensed via global average pooling:

$$\mathbf{z}_H = \mathrm{GAP}(\mathbf{C}_{\mathrm{HS}}), \tag{12}$$
$$\mathbf{z}_S = \mathrm{GAP}(\mathbf{C}_{\mathrm{SH}}), \tag{13}$$

and subsequently fused:

$$\mathbf{z}_t = [\mathbf{z}_H \oplus \mathbf{z}_S], \tag{14}$$

where $\oplus$ denotes the concatenation operator.

6) PREDICTION LAYER

The fused vector $\mathbf{z}_t$ is passed through a dense layer with ReLU activation and dropout regularization:

$$\mathbf{k}_t = \mathrm{Dropout}(\mathrm{ReLU}(\mathrm{Dense}(\mathbf{z}_t))), \tag{15}$$

followed by a sigmoid-activated output unit:

$$\hat{y}_{t+1} = \sigma(\mathrm{Dense}(\mathbf{k}_t)), \tag{16}$$

which yields the probability of an upward price movement on day $t + 1$.

7) TRAINING STRATEGY

The model is trained using the binary cross-entropy loss:

$$\mathcal{L} = -\left[y_{t+1}\log\hat{y}_{t+1} + (1 - y_{t+1})\log(1 - \hat{y}_{t+1})\right], \tag{17}$$

with class weighting to address label imbalance. Dropout and early stopping are used to prevent overfitting.

8) DESIGN RATIONALE

This architecture empowers the model to learn hierarchical representations within each modality while dynamically capturing interdependencies across financial signals and sentiment narratives. The use of bidirectional attention between modalities represents a key innovation, enabling context-aware fusion critical for volatile and sentiment-sensitive domains like cryptocurrency forecasting.

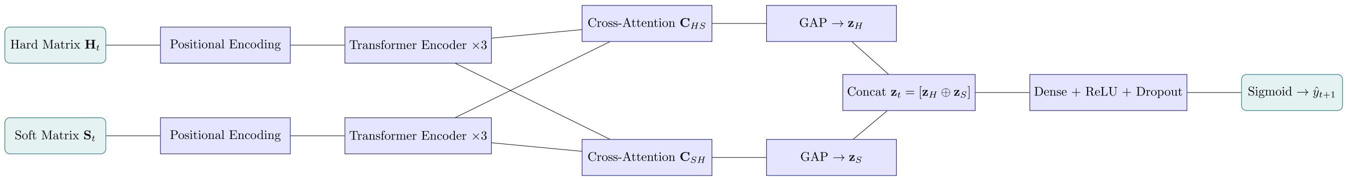

FIGURE 3. Illustration of the proposed dual–stream Transformer architecture with bidirectional cross–attention fusion for multimodal cryptocurrency trend prediction.

TABLE 3. Temporal and statistical distribution of Bitcoin dataset.

Subset	Start Date	End Date	Tweet Count	Days	Up Days	Down Days
Training	2015-04-06	2020-09-03	33,023	1,978	1,083	895
Validation	2020-09-04	2021-11-01	35,797	424	235	189
Testing	2021-11-02	2022-12-31	35,476	425	198	227

III. EXPERIMENTS AND RESULTS

To rigorously assess the empirical efficacy of the proposed cross-attentive dual-stream Transformer framework, a multifaceted experimental protocol is employed. This encompasses not only standard classification metrics but also comprehensive financial backtesting to gauge real-world investment viability. Comparative baselines include both state-of-the-art LSTM-based architectures and transformer-driven multimodal fusion variants.

A. DATASET CONSTRUCTION AND PREPROCESSING

Bitcoin, as the most capitalized and actively traded cryptocurrency, is selected as the target asset for evaluation. The experimental dataset is composed of both quantitative financial data (hard data) and qualitative sentiment data (soft data), enabling a multimodal learning process.

a: QUANTITATIVE FEATURES (HARD INFORMATION)

Historical Bitcoin trading data was retrieved from Yahoo Finance, spanning from 6 April 2015 to 31 December 2022. This raw data was preprocessed into 52 technical indicators encompassing momentum oscillators, volatility measures, volume-based metrics, and trend-following features. Redundancy was mitigated through Pearson correlation analysis ($r > 0.95$), reducing the dimensionality to 36. Subsequently, SHAP-based interpretability was leveraged to retain the top 30 features with the highest explanatory power for price movement classification.

b: TEXTUAL FEATURES (SOFT INFORMATION)

A high-fidelity soft dataset was curated from the X platform by monitoring 12 cryptocurrency-focused accounts. Post quality-control filtering, 8 accounts remained, contributing over $104,000$ tweets. FinBERT, a domain-optimized BERT variant, was used to extract sentiment vectors (positive, neutral, negative) at the tweet level. These vectors were then aggregated on a daily basis to yield structured soft inputs $\mathbf{S}_t \in \mathbb{R}^{T\times 3}$, aligned with hard features.

The distribution of Bitcoin-related news and corresponding market labels across the training, validation, and testing splits is systematically detailed in Table 3, providing a comprehensive overview of the temporal span and class balance underpinning the experimental evaluation.

While the present study focuses on Bitcoin, this decision was based on its dominant market capitalization, deep liquidity, and the uniquely high availability of both structured market indicators and reliable, temporally aligned sentiment signals. Other major cryptocurrencies exhibited significantly lower volumes of high-quality sentiment data during the same period, limiting their suitability for robust multimodal modeling. Nonetheless, the HSIF framework is inherently modular and asset-agnostic. The hard information stream can be readily adapted to any asset with sufficient historical trading data, and the soft sentiment stream can be reconstructed using token-specific or platform-wide social and news discourse. Crucially, the bidirectional cross-attention mechanism is not asset-dependent and can dynamically learn inter-modal dependencies across varying market contexts.

B. EXPERIMENTAL PROTOCOL

The Transformer-based architecture was implemented using a DL framework. A sequence length of $T = 60$ was chosen to capture mid-term market memory. Data normalization was performed via min-max scaling to ensure feature homogeneity. The dataset was split chronologically into training (70%), validation (15%), and testing (15%) sets to preserve temporal integrity.

Hyperparameters were exhaustively tuned via grid search. The final model employs three Transformer encoder layers, each with four multi-head attention mechanisms and 64-dimensional key vectors per head. Dropout regularization and early stopping were utilized to curb overfitting. Table 4 details the selected configurations.

Class imbalance was addressed using weighted cross-entropy loss. Additionally, decision thresholds were optimized on the validation set via F1-maximization, scanning thresholds from 0.30 to 0.70.

TABLE 4. Hyperparameter configuration for Transformer architecture.

Parameter	Value
Batch Size	32
Epochs	50
Optimizer	Adam
Learning Rate	0.001
Transformer Layers	3
Attention Heads	4
Head Dimension	64
Dropout (Fusion Layer)	0.30
Dropout (Dense Layer)	0.20
Runs per Model	20

To determine the optimal architectural depth for the transformer encoder, a series of ablation studies were conducted by varying the number of encoder layers from one to four. As presented in Table 5, a configuration with three stacked transformer layers achieved the highest overall classification performance.

C. EVALUATION FRAMEWORK

To ensure a rigorous and multidimensional assessment of the proposed Transformer-based prediction framework, both classification and financial performance metrics are employed. This dual-perspective evaluation strategy enables not only a prediction assessment of the model, but also its practical efficacy in real-world trading scenarios.

1) CLASSIFICATION METRICS

Three primary classification metrics are utilized to measure the model's performance in binary trend prediction: Accuracy, F1 score, and Matthews Correlation Coefficient (MCC). These metrics collectively capture the balance between true predictions and class distribution, providing a holistic view of classification quality.

- **Accuracy**: Denotes the proportion of correctly classified instances among the total number of samples. It is computed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{18}$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

- **F1 Score**: A harmonic mean of precision and recall, particularly valuable in imbalanced classification problems. It quantifies the model's ability to correctly predict the minority class without being biased by the dominant one:

$$\text{F1 Score} = \frac{2 \times TP}{2 \times TP + FP + FN}, \tag{19}$$

- **MCC**: A robust and informative metric that evaluates classification performance across all four components of the confusion matrix. Unlike accuracy, MCC remains reliable under severe class imbalance and is defined as:

$$\text{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}, \tag{20}$$

MCC values range from -1 (complete disagreement) to $+1$ (perfect prediction), with 0 indicating random guessing.

2) FINANCIAL PERFORMANCE METRICS

To quantify the economic viability of the proposed framework, the model is evaluated in a simulated trading environment. This simulation employs model predictions to drive buy/sell decisions under realistic market conditions, including transaction costs and capital constraints. Three widely accepted financial indicators are adopted:

- **Final Capital (TMoney)**: Represents the accumulated capital at the end of the trading simulation, beginning with an initial capital of 100,000 USD. Trading decisions are governed by Algorithm 1, with a fixed transaction commission of 0.1% applied to each operation.
- **Return Percentage (R)**: Measures the total percentage gain or loss relative to the initial investment, offering a direct assessment of profitability:

$$\text{Return (\%)} = \frac{TMoney_{\text{final}} - TMoney_{\text{initial}}}{TMoney_{\text{initial}}} \times 100, \tag{21}$$

where $TMoney_{\text{initial}}$ and $TMoney_{\text{final}}$ denote the starting and final capital, respectively.

- **Sharpe Ratio (SR)**: A critical risk-adjusted performance measure, the Sharpe ratio quantifies the excess return per unit of volatility. It is calculated as:

$$\text{Sharpe Ratio} = \frac{\mu_R}{\sigma_R}, \tag{22}$$

where μ_R denotes the mean of daily returns and σ_R represents the standard deviation of daily returns. A higher Sharpe ratio implies a more favorable trade-off between return and risk.

This dual-metric evaluation framework, spanning both classification integrity and financial profitability, provides a robust lens through which the real-world readiness of the proposed model can be rigorously examined.

D. COMPARATIVE RESULTS AND INSIGHTS

Table 6 presents a comprehensive evaluation of seven model configurations, encompassing hard only, simple concatenation, and advanced fusion strategies across three sentiment encoders: CryptoBERT, RoBERTa, and FinBERT. Notably, the fusion-based architecture leveraging FinBERT achieves the highest predictive performance, delivering an accuracy of 97.48%, F1 score of 97.27%, MCC of 94.98%, return percentage of 26.64%, and Sharpe ratio of 0.3344, significantly

TABLE 5. **Comparison of model performance with different numbers of Transformer layers (mean ± standard deviation).**

Number of Layers	Accuracy	F1 Score	MCC	Return (%)	Sharpe Ratio
1	0.938 ± 0.0113	0.945 ± 0.00757	0.882 ± 0.0211	$\mathbf{25.39 \pm 0.613}$	0.317 ± 0.0075
2	0.940 ± 0.0113	0.945 ± 0.00764	0.884 ± 0.0212	25.03 ± 0.604	0.315 ± 0.0074
3	$\mathbf{0.943 \pm 0.0112}$	$\mathbf{0.949 \pm 0.00746}$	$\mathbf{0.893 \pm 0.0194}$	25.31 ± 0.596	$\mathbf{0.319 \pm 0.0063}$
4	0.933 ± 0.0112	0.937 ± 0.00750	0.871 ± 0.0208	25.17 ± 0.607	0.315 ± 0.0075

Algorithm 1 Cryptocurrency Trading Algorithm With Commission, Used to Simulate Trading Decisions Based on Predicted Market Trends

Input: Prediction Results, Opening and Closing Prices
Output: Final Capital (TMoney)

$TMoney \leftarrow 100,000$
$label\_previous \leftarrow -1$
$Commission\_Rate \leftarrow 0.001$
while Cryptocurrency Price Prediction Data Available **do**
  $label \leftarrow$ Prediction Result
  if $label \neq label\_previous$ **then**
    if $label = 1$ **then**
      $Commission \leftarrow Commission\_Rate \times TMoney$
      Buy Cryptocurrency
      $TMoney \leftarrow TMoney + Closing\_Price \times Number\_of\_Stocks - Commission$
    else
      Sell Cryptocurrency
      $Commission \leftarrow Commission\_Rate \times TMoney$
      $TMoney \leftarrow TMoney + Closing\_Price \times Number\_of\_Stocks - Commission$
    end if
  end if
  $label\_previous \leftarrow label$
end while
return $TMoney$

outperforming its simple concatenation counterpart and all other baselines.

The empirical findings highlight the importance of modality interaction: while simple concatenation preserves feature representation from each stream, it lacks the contextual synergy enabled by cross-attention. By contrast, the fusion strategy explicitly models interdependencies between quantitative indicators and sentiment signals, thus capturing a richer joint feature space.

These results demonstrate that domain-specific sentiment analysis, when properly fused with structured quantitative data through a Transformer-based dual-stream cross-attention mechanism, offers a significant edge in cryptocurrency trend prediction. In particular, the superiority of the FinBERT-based fusion model underscores the importance of financial language modeling in capturing market-relevant sentiment nuances.

E. BENCHMARKING AGAINST SEQUENTIAL MODELS

To underscore the superiority of attention mechanisms over recurrence-based models, comparative experiments are conducted with Bidirectional LSTM (BiLSTM), Stacked LSTM, and CNN-LSTM architectures. Although BiLSTM attains strong classification scores, all recurrent architectures fail to produce positive financial returns, suggesting overfitting to short-term temporal signals and an inability to adapt to market non-stationarity.

To underscore the superiority of attention mechanisms over recurrence-based models, comparative experiments are conducted with BiLSTM, Stacked LSTM, and CNN-LSTM architectures. As shown in Table 7, although BiLSTM achieves strong classification performance, all recurrent architectures fail to produce positive financial returns, suggesting overfitting to short-term temporal signals and an inability to adapt to market non-stationarity.

The findings conclusively demonstrate that integrating domain-specific sentiment embeddings with quantitative indicators through a bidirectional cross-attention Transformer architecture yields substantial gains in both predictive accuracy and financial return. The model not only excels at capturing nonlinear dependencies across modalities, but also exhibits resilience in volatile market regimes, an essential trait for deployment in algorithmic trading pipelines.

F. CASE STUDY: MODEL BEHAVIOR DURING SIGNIFICANT MARKET EVENT

To further evaluate the interpretability of the HSIF model and understand how it incorporates sentiment signals, we examine its prediction during a major market event driven by public sentiment. On 15 October 2021, the US Securities and Exchange Commission (SEC) approved the first Bitcoin futures Exchange-Traded Fund (ETF), triggering a surge in positive market sentiment. FinBERT-derived sentiment scores for that day were strongly optimistic (positive $= 0.95$, neutral $= 0.045$, negative $= 0.004$). HSIF correctly predicted an upward movement ($\hat{y}_{t+1} = 1$), which aligned with the actual market outcome. This finding highlights the ability of the model to incorporate qualitative investor sentiment in a meaningful and interpretable way.

IV. DISCUSSION

The empirical findings of this study underscore the transformative potential of multimodal DL and domain-specific NLP in enhancing cryptocurrency trend prediction. The proposed

TABLE 6. Performance comparison across fusion architectures and sentiment encoders (mean ± standard deviation).

Model Configuration	Accuracy	F1 score	MCC	Return (%)	SR
Hard Only	0.9427 ± 0.0112	0.9493 ± 0.0076	0.8932 ± 0.0211	25.31 ± 0.60	0.3191 ± 0.0063
Hard + CryptoBERT (Simple Concatenation)	0.9458 ± 0.0112	0.9527 ± 0.0174	0.8951 ± 0.0418	25.58 ± 0.69	0.3226 ± 0.0047
Hard + RoBERTa (Simple Concatenation)	0.9462 ± 0.0143	0.9532 ± 0.0207	0.8947 ± 0.0378	25.80 ± 0.76	0.3250 ± 0.0052
Hard + FinBERT (Simple Concatenation)	0.9465 ± 0.0113	0.9533 ± 0.0197	0.8966 ± 0.0356	25.89 ± 0.80	0.3256 ± 0.0055
Hard + CryptoBERT (Fusion)	0.9695 ± 0.0096	0.9671 ± 0.0109	0.9395 ± 0.0187	26.57 ± 0.85	0.3333 ± 0.0064
Hard + RoBERTa (Fusion)	0.9659 ± 0.0249	0.9643 ± 0.0237	0.9333 ± 0.0445	26.60 ± 0.87	0.3335 ± 0.0071
Hard + FinBERT (Fusion)	**0.9748 ± 0.0106**	**0.9727 ± 0.0117**	**0.9498 ± 0.0207**	**26.64 ± 0.98**	**0.3344 ± 0.0079**

TABLE 7. Performance comparison between the proposed HSIF model and LSTM-based architectures, with all results reported as mean ± standard deviation.

Model	Accuracy	F1 score	MCC	Return (%)	SR
BiLSTM	0.9666 ± 0.0090	0.9652 ± 0.0090	0.9346 ± 0.0169	-18.08 ± 1.08	-0.2106 ± 0.0106
Stacked LSTM	0.9656 ± 0.0079	0.9640 ± 0.0079	0.9322 ± 0.0149	-18.19 ± 0.55	-0.2151 ± 0.0091
CNN-LSTM	0.9221 ± 0.0147	0.9171 ± 0.0151	0.8441 ± 0.0290	-22.57 ± 1.26	-0.2638 ± 0.0135
Proposed Model	**0.9748 ± 0.0106**	**0.9727 ± 0.0117**	**0.9498 ± 0.0207**	**26.64 ± 0.98**	**0.3344 ± 0.0079**

HSIF framework, underpinned by a Transformer-based architecture and bidirectional cross-attention, significantly outperforms existing benchmarks across both classification and financial metrics.

A. ARCHITECTURAL INSIGHTS AND CORE CONTRIBUTIONS

A primary innovation of this work lies in its ability to dynamically integrate heterogeneous data sources, structured numerical indicators and unstructured textual sentiment, using cross-modal attention mechanisms. While traditional models are limited to unimodal inputs or simple concatenation, the proposed architecture captures intricate interdependencies between market signals and investor sentiment in a temporally consistent fashion. This fusion not only improves prediction fidelity but also introduces interpretability into the decision-making pipeline.

The architecture reflects a deeper conceptual shift from conventional data fusion toward information fusion. Data fusion techniques, which integrate disparate sources of raw or semi-structured data, enhance the reliability, accuracy, and consistency of predictions by mitigating issues such as data uncertainty, incompleteness, and inconsistency [29]. In contrast, information fusion operates at a higher abstraction level, combining semantically processed, structured, and context-aware features from multiple modalities to support deeper insight extraction and informed decision-making [30]. The proposed HSIF model embodies this distinction by not only merging raw features but also interpreting them through domain-specific sentiment encoding and attention-driven integration, ultimately achieving more context-sensitive and actionable outputs.

The superior performance of the FinBERT-driven sentiment stream further illustrates the critical role of domain-adapted language models in financial forecasting. Unlike general-purpose models, FinBERT exhibits enhanced semantic precision in identifying market-relevant cues, particularly in the context of emotionally charged and speculation-driven markets like cryptocurrency. This aligns with emerging trends in finance-specific language modeling, where pretrained transformers tailored to economic contexts show clear empirical benefits.

Table 8 presents a comprehensive comparison between HSIF and a range of recent Bitcoin forecasting models, including other multimodal and DL-based approaches. Although several previous works have integrated sentiment data or applied DL techniques, HSIF introduces a uniquely cohesive and technically novel design.

To highlight the architectural innovations of HSIF, we compare it with two recent multimodal models, CBITS [39] and PreBit [38]. Unlike CBITS, which uses static sentiment scores and TabNet without temporal modeling or joint optimization, HSIF incorporates dynamic bidirectional cross-attention and trains both data streams end-to-end. Similarly, PreBit relies on the simple concatenation of CNN-based market characteristics and FinBERT tweet embeddings but lacks temporal fusion and focuses only on extreme price movement prediction. In contrast, HSIF offers a more expressive and adaptable framework by modeling interactions between sentiment and market indicators over time, enhancing both forecast accuracy and practical applicability.

HSIF incorporates SHAP-based feature filtering, supports joint end-to-end training, and is evaluated using investment-grade financial metrics such as return and Sharpe ratio. These features make HSIF a robust and practical tool for real-world trading strategies.

This expanded comparison demonstrates that HSIF is not only a high-performing model but also an architecturally distinct and finance-aware approach, combining interpretability, modularity, and scalability for multimodal trend prediction.

TABLE 8. A comparative analysis of the proposed method versus existing approaches.

Work	Dataset	Objective	Methods	Performance
2023 [37]	On-chain Bitcoin data from CryptoQuant + whale alert tweets	Classification Bitcoin trend and volatility prediction	Reinforcement Learning (RL) Q-Learning with Markov Decision Process	Accuracy = 0.90 F1 Score = 0.847
2023 [10]	CoinMarketCap price data + asset/interest-based data + Google Trends	Classification Price trend	LSTM + ReGAT (Graph Attention over Crypto Network)	Accuracy = 0.6297 F1 Score = 0.5469
2023 [38]	Twitter data + Technical indicators	Classification Bitcoin price movement	PreBit: Dual-branch Fusion CNN (technical) + FinBERT (tweets) Fusion via Concatenation	Up Accuracy = 0.737 Down Accuracy = 0.8137
2023 [39]	Multimodal dataset: Crypto news (Korean) + Binance BTC/USDT chart data	Classification Bitcoin trading actions	CBITS: TabNet + CryptoRoBERTa Sentiment Scores Feature-level Fusion	Accuracy = 0.7335 F1 Score = 0.6983
2024 [40]	Historical price data + Technical indicators	Classification Bitcoin trend prediction	Attention-LSTM	Accuracy = 0.7384 F1 Score = 0.6235
2024 [41]	Historical Price Data + News and Social Media Signals	Prediction of Bitcoin Price Movement and Trading Strategy Development	LSTM Neural Network + Multi-source Signal Integration + Financial Data Fusion	Accuracy = 0.823 Sharpe Ratio = 1.27
2024 [42]	Historical price data + Technical indicators	Classification Multi-asset crypto trend prediction	Attention-based CNN–LSTM	Buy-Hold ER = 6.05% Long-Short ER = 7.72%
2025 [43]	Technical indicators + Twitter posts	Classification Bitcoin price range prediction	CART Decision Tree + Twitter-roBERTa Sentiment Analysis	Accuracy = 0.62
This study	Historical price data + Technical indicators + X (Twitter) Bitcoin news	Classification Bitcoin trend prediction	SHAP-based Feature Filtering + HSIF: Transformer-based Dual-Stream Model + Bidirectional Cross-Attention Fusion	Accuracy = 0.9748 F1 Score = 0.9727 MCC = 0.9498 Return = 26.64% Sharpe Ratio = 0.3344

Compared to state-of-the-art baselines (see Table 8), the HSIF model achieves a substantial leap in predictive accuracy (97.48%) and profitability (26.64% return with a Sharpe ratio of 0.3344), suggesting that cross-attentive fusion of soft and hard data yields both algorithmic alpha, excess return driven by the predictive edge of model, and improved market adaptivity. This finding challenges the dominance of traditional LSTM-based sequential models, which, despite high classification scores, often fail to generalize in volatile and nonstationary trading environments.

Furthermore, the dual-stream design promotes temporal abstraction and modality disentanglement, allowing the model to isolate and recombine relevant temporal signals across data types. This contrasts with conventional attention or CNN-LSTM hybrids, which are prone to overfitting short-term trends without capturing broader macrobehavior or sentiment drifts.

From a practical standpoint, our results advocate the integration of social signals into automated trading frameworks. In particular, the robustness of the architecture across diverse time frames and news cycles suggests high viability for deployment in live trading environments, potentially improving risk-adjusted returns and market timing decisions.

B. ETHICAL AND PRACTICAL CONSIDERATIONS

This study uses sentiment data exclusively from a curated set of reliable, domain-specific X accounts focused on Bitcoin, rather than from random users. This reduces exposure to misinformation, noise, and manipulative content. By restricting the dataset to high-credibility sources and aggregating sentiment signals at the daily level, the need for separate mechanisms for the detection of fake news was effectively mitigated. All tweets were publicly available, and no personal or private data was collected. Nevertheless, we acknowledge that sentiment-based models may still be vulnerable to manipulation. Therefore, any real-world implementation of HSIF should include safeguards such as human oversight, anomaly detection, and risk controls to ensure responsible use.

C. ADVANCING FRAMEWORK: FUTURE DIRECTIONS

Despite the promising results achieved by the proposed model, there remain several challenges in the field of cryptocurrency price prediction that warrant exploration in future research. While this study focuses on Bitcoin due to its dominant market capitalization and the availability of high-quality sentiment and market data, the proposed HSIF framework is inherently modular and asset-agnostic. Its design supports extension to other cryptocurrencies or financial instruments, as well as integration of new data sources and deployment-aware modeling enhancements. Key directions for future work include:

- **Fake News Detection:** In this work, sentiment signals are derived exclusively from verified and reputable sources to ensure data reliability. To extend the framework to broader social media and news platforms, where misinformation is more prevalent, future efforts will incorporate robust fake news detection mechanisms.

These may include real-time fact-checking systems, credibility scoring, and anomaly detection algorithms to filter unreliable content, which is critical for safe and ethical deployment in automated trading systems.
- **Prioritization of News:** Not all news carries equal market impact. A future direction involves developing relevance-aware attention modules or external ranking systems to assign weights to news items based on contextual importance and historical market influence.
- **Multi-Cryptocurrency Forecasting:** While this study focuses on Bitcoin due to data quality and market dominance, the HSIF framework is inherently modular and asset-agnostic. Future work will extend it to other major cryptocurrencies, enabling portfolio-level forecasting and cross-asset sentiment modeling.
- **Incorporating Realistic Trading Constraints:** Future work will also explore integrating realistic trading constraints, such as slippage, volatility, and transaction costs, into the model's trading algorithm. This will help to simulate real world conditions more accurately and improve the practical applicability of the model in live trading environments.

In summary, this study not only advances the state-of-the-art in crypto-financial forecasting but also lays a foundation for broader applications of multimodal learning and sentiment-aware modeling in high-volatility domains.

V. CONCLUSION

This study introduces HSIF, a novel Transformer-based dual-stream model that fuses market data and sentiment through bidirectional cross attention for cryptocurrency trend prediction. Our results show state-of-the-art accuracy (97.48%) and financial returns (26.64%), demonstrating the ability of HSIF to integrate structured and unstructured data in volatile markets. Using FinBERT for sentiment analysis and SHAP for feature selection, HSIF enhances the fusion of market sentiment with technical analysis. This research paves the way for future work in cross-asset forecasting and live trading, supporting the development of autonomous sentiment-aware trading systems.

REFERENCES

[1] Z. Zhou, Z. Song, H. Xiao, and T. Ren, "Multi-source data driven cryptocurrency price movement prediction and portfolio optimization," *Expert Syst. Appl.*, vol. 219, Jun. 2023, Art. no. 119600.

[2] Y. Xiang, Y. Lei, D. Bao, T. Li, Q. Yang, W. Liu, W. Ren, and K.-K.-R. Choo, "BABD: A Bitcoin address behavior dataset for pattern analysis," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 2171–2185, 2024.

[3] S. Squarepants, "Bitcoin: A peer-to-peer electronic cash system," 2008. [Online]. Available: https://bitcoin.org/en/bitcoin-paper

[4] A. Wahrstätter, A. Taudes, and D. Svetinovic, "Reducing privacy of CoinJoin transactions: Quantitative Bitcoin network analysis," *IEEE Trans. Dependable Secure Comput.*, vol. 21, no. 5, pp. 4543–4558, Sep. 2024.

[5] A. F. M. S. Shah, M. A. Karabulut, A. F. M. S. Akhter, N. Mustari, A.-S.-K. Pathan, K. M. Rabie, and T. Shongwe, "On the vital aspects and characteristics of cryptocurrency—A survey," *IEEE Access*, vol. 11, pp. 9451–9468, 2023.

[6] X. Dong and M. Yu, "Time-varying effects of macro shocks on cross-border capital flows in China's bond market," *Int. Rev. Econ. Finance*, vol. 96, Nov. 2024, Art. no. 103720.

[7] I. Gurrib and F. Kamalov, "Predicting Bitcoin price movements using sentiment analysis: A machine learning approach," *Stud. Econ. Finance*, vol. 39, no. 3, pp. 347–364, Apr. 2022.

[8] M. Alizadeh, D. S. Zadeh, B. Moshiri, and A. Montazeri, "Development of a customer churn model for banking industry based on hard and soft data fusion," *IEEE Access*, vol. 11, pp. 29759–29768, 2023.

[9] S. Goutte, H.-V. Le, F. Liu, and H.-J. von Mettenheim, "Deep learning and technical analysis in cryptocurrency market," *Finance Res. Lett.*, vol. 54, Jun. 2023, Art. no. 103809.

[10] C. Zhong, W. Du, W. Xu, Q. Huang, Y. Zhao, and M. Wang, "LSTM-ReGAT: A network-centric approach for cryptocurrency price trend prediction," *Decis. Support Syst.*, vol. 169, Jun. 2023, Art. no. 113955.

[11] N. Majidi, M. Shamsi, and F. Marvasti, "Algorithmic trading using continuous action space deep reinforcement learning," *Expert Syst. Appl.*, vol. 235, Jan. 2024, Art. no. 121245.

[12] M. Qureshi, H. Iftikhar, P. C. Rodrigues, M. Z. Rehman, and S. A. A. Salar, "Statistical modeling to improve time series forecasting using machine learning, time series, and hybrid models: A case study of Bitcoin price forecasting," *Mathematics*, vol. 12, no. 23, p. 3666, Nov. 2024.

[13] K. A. Coulter, "The impact of news media on Bitcoin prices: Modelling data driven discourses in the crypto-economy with natural language processing," *Roy. Soc. Open Sci.*, vol. 9, no. 4, Apr. 2022, Art. no. 220276.

[14] S. M. Dashtaki, M. Alizadeh, and B. Moshiri, "Stock market prediction using hard and soft data fusion," in *Proc. 13th Int. Conf. Inf. Knowl. Technol. (IKT)*, Dec. 2022, pp. 1–7.

[15] Y. Ma, R. Mao, Q. Lin, P. Wu, and E. Cambria, "Multi-source aggregated classification for stock price movement prediction," *Inf. Fusion*, vol. 91, pp. 515–528, Mar. 2023.

[16] L. Jing, X. Fan, D. Feng, C. Lu, and S. Jiang, "A patent text-based product conceptual design decision-making approach considering the fusion of incomplete evaluation semantic and scheme beliefs," *Appl. Soft Comput.*, vol. 157, May 2024, Art. no. 111492.

[17] F. Feizian and B. Amiri, "Cryptocurrency price prediction model based on sentiment analysis and social influence," *IEEE Access*, vol. 11, pp. 142177–142195, 2023.

[18] S. Otabek and J. Choi, "From prediction to profit: A comprehensive review of cryptocurrency trading strategies and price forecasting techniques," *IEEE Access*, vol. 12, pp. 87039–87064, 2024.

[19] M. N. Ashtiani and B. Raahemi, "News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review," *Expert Syst. Appl.*, vol. 217, May 2023, Art. no. 119509.

[20] A. Ibrahim, R. Kashef, and L. Corrigan, "Predicting market movement direction for Bitcoin: A comparison of time series modeling methods," *Comput. Electr. Eng.*, vol. 89, Jan. 2021, Art. no. 106905.

[21] N. P. Patel, R. Parekh, N. Thakkar, R. Gupta, S. Tanwar, G. Sharma, I. E. Davidson, and R. Sharma, "Fusion in cryptocurrency price prediction: A decade survey on recent advancements, architecture, and potential future directions," *IEEE Access*, vol. 10, pp. 34511–34538, 2022.

[22] M. Hosseini Chagahi, N. Delfan, S. Mohammadi Dashtaki, B. Moshiri, and M. Jalil Piran, "Explainable AI for fraud detection: An attention-based ensemble of CNNs, GNNs, and a confidence-driven gating mechanism," 2024, *arXiv:2410.09069*.

[23] S. K. Sahu, A. Mokhade, and N. D. Bokde, "An overview of machine learning, deep learning, and reinforcement learning-based techniques in quantitative finance: Recent progress and challenges," *Appl. Sci.*, vol. 13, no. 3, p. 1956, Feb. 2023.

[24] G. Bansal, V. Chamola, G. Kaddoum, M. J. Piran, and M. Alrashoud, "Next generation stock exchange: Recurrent neural learning model for distributed ledger transactions," *Comput. Netw.*, vol. 193, Jul. 2021, Art. no. 107998.

[25] S. Islam, H. Elmekki, A. Elsebai, J. Bentahar, N. Drawel, G. Rjoub, and W. Pedrycz, "A comprehensive survey on applications of transformers for deep learning tasks," *Expert Syst. Appl.*, vol. 241, May 2024, Art. no. 122666.

[26] 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. 121846–121863, 2024.

[27] A. Iranfar, M. Soleymannejad, B. Moshiri, and H. D. Taghirad, "Natural language processing and soft data for motor skill assessment: A case study in surgical training simulations," *Comput. Methods Programs Biomed.*, vol. 264, Jun. 2025, Art. no. 108686.

[28] M. A. Izadi and E. Hajizadeh, ''Time series prediction for cryptocurrency markets with transformer and parallel convolutional neural networks,'' *Appl. Soft Comput.*, vol. 177, Jun. 2025, Art. no. 113229.
[29] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, ''A survey on machine learning for data fusion,'' *Inf. Fusion*, vol. 57, pp. 115–129, May 2020.
[30] M. A. Becerra, C. Tobón, A. E. Castro-Ospina, and D. H. Peluffo-Ordóñez, ''Information quality assessment for data fusion systems,'' *Data*, vol. 6, no. 6, p. 60, Jun. 2021.
[31] Twitter. *Twitter Platform for Social Media Content*. Accessed: Mar. 1, 2023. [Online]. Available: https://twitter.com/
[32] D. Araci, ''FinBERT: Financial sentiment analysis with pre-trained language models,'' 2019, *arXiv:1908.10063*.
[33] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, ''RoBERTa: A robustly optimized BERT pretraining approach,'' 2019, *arXiv:1907.11692*.
[34] M. Kulakowski and F. Frasincar, ''Sentiment classification of cryptocurrency-related social media posts,'' *IEEE Intell. Syst.*, vol. 38, no. 4, pp. 5–9, Jul. 2023.
[35] Yahoo Finance. *Yahoo Finance: Historical Market Data*. Accessed: Mar. 1, 2023. [Online]. Available: https://finance.yahoo.com/
[36] R. M. Dashtaki, S. M. Dashtaki, E. Heydari-Bafrooei, and M. J. Piran, ''Enhancing the predictive performance of molecularly imprinted polymer-based electrochemical sensors using a stacking regressor ensemble of machine learning models,'' *ACS Sensors*, vol. 10, no. 4, pp. 3123–3133, Apr. 2025.
[37] M. Azamjon, O. Sattarov, and J. Cho, ''Forecasting Bitcoin volatility through on-chain and whale-alert tweet analysis using the Q-learning algorithm,'' *IEEE Access*, vol. 11, pp. 108092–108103, 2023.
[38] Y. Zou and D. Herremans, ''PreBit—A multimodal model with Twitter FinBERT embeddings for extreme price movement prediction of Bitcoin,'' *Expert Syst. Appl.*, vol. 233, Dec. 2023, Art. no. 120838.
[39] G. Kim, M. Kim, B. Kim, and H. Lim, ''CBITS: Crypto BERT incorporated trading system,'' *IEEE Access*, vol. 11, pp. 6912–6921, 2023.
[40] M.-C. Lee, ''Bitcoin trend prediction with attention-based deep learning models and technical indicators,'' *Systems*, vol. 12, no. 11, p. 498, Nov. 2024.
[41] H. Subramanian, P. Angle, F. Rouxelin, and Z. Zhang, ''A decision support system using signals from social media and news to predict cryptocurrency prices,'' *Decis. Support Syst.*, vol. 178, Mar. 2024, Art. no. 114129.
[42] P. Peng, Y. Chen, W. Lin, and J. Z. Wang, ''Attention-based CNN–LSTM for high-frequency multiple cryptocurrency trend prediction,'' *Expert Syst. Appl.*, vol. 237, Mar. 2024, Art. no. 121520.
[43] L. Shang, ''Sentiment-driven Bitcoin price range forecasting: Enhancing CART decision trees with high-dimensional indicators and Twitter dynamics,'' *IEEE Access*, vol. 13, pp. 60508–60518, 2025.

SAEED MOHAMMADI DASHTAKI received the B.Sc. degree in electrical engineering from Isfahan University of Technology, Isfahan, Iran, in 2019, and the M.Sc. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2023. He has been a Research Assistant and a Teaching Assistant with the University of Tehran and also a Research Assistant with Sejong University, Seoul, South Korea, where his research contributions at both universities have focused on machine learning, data/information fusion, neural networks and deep learning, multi-modal learning, trustworthy AI, financial engineering, and healthcare. He has published several articles in high-impact journals, such as ''IEEE TRANSACTIONS journals.'' His research interests include artificial intelligence, deep learning, data/information fusion, financial engineering, medical image analysis, robotics, and control theory.

MEHDI HOSSEINI CHAGAHI received the B.Sc. degree in electrical and computer engineering (Hons.) and the M.Sc. degree in electrical and computer engineering from the University of Tehran, Tehran, Iran, one of the top-ranked universities in the country. He was ranked 56th in the nationwide master's entrance examination. His research contributions include publications in IEEE Transactions journals, along with service as a reviewer for leading international journals. He has been a Research Assistant with Sejong University, Seoul, South Korea, where he worked on projects in cardiovascular disease prediction, multi-modal learning for medical diagnosis, and AI-based brain hemorrhage detection. He also served as a Teaching Assistant with the University of Tehran, where he guided students in machine learning and deep learning. In addition, he leads multiple research teams in the field of artificial intelligence. For his M.Sc. thesis, he developed a fuzzy-based vision transformer for brain hemorrhage diagnosis from CT scans. The model was successfully deployed in a hospital setting, where it significantly improved both diagnostic speed and accuracy. His research interests include artificial intelligence in healthcare, trustworthy AI, fuzzy logic, and mobile computing.

AEIN BAHADORI received the B.Sc. degree in electrical engineering (control systems) from Isfahan University of Technology, Iran, in 2020, and the M.Sc. degree in control system engineering (with a specialization in machine learning) from the University of Padova, Italy, in 2024. His research interests include data science, deep learning, machine learning, computer vision, robotics, and control theory.

BEHZAD MOSHIRI (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in control systems engineering from the University of Manchester Institute of Science and Technology (UMIST), in 1987 and 1991, respectively. He has been an Adjunct Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, since 2014. He is currently a Full Professor of control systems engineering with the School of Electrical and Computer Engineering, University of Tehran, Iran. He is the author/co-author of more than 360 articles, including more than 130 journal articles and more than 20 book chapters. His research interests include advanced industrial control, advanced instrumentation systems, data fusion theory, and its applications in areas such as robotics, process control, mechatronics, information technology (IT), intelligent transportation systems (ITS), bioinformatics, and financial engineering. He has been a member of the International Society of Information Fusion (ISIF), since 2002, and a member of the Waterloo AI Institute, since 2019. He has also been serving as the Chairperson for the IEEE Control Systems Chapter in the IEEE Iran Section, since December 2018.

MD. JALIL PIRAN (Senior Member, IEEE) received the Ph.D. degree in electronics and information engineering from Kyung Hee University, South Korea, in 2016. He was a Postdoctoral Fellow with Kyung Hee University, until 2017. He holds a distinguished academic background and is currently an Associate Professor with the Department of Computer Science and Engineering, Sejong University, Seoul, South Korea. He has made significant contributions to the field of artificial intelligence and data science through his extensive research publications in esteemed international journals and conferences. His areas of expertise encompass machine learning, data science, big data, the Internet of Things (IoT), and cyber security. In addition to his research endeavors, he actively engages with scholarly journals as an Editor, including IEEE Transactions on Intelligent Transportation Systems, *Engineering Applications of Artificial Intelligence* (Elsevier), *Physical Communication* (Elsevier), and *Computer Communications* (Elsevier). He represents South Korea as an Active Delegate to the Moving Picture Experts Group (MPEG). His outstanding research contributions have been recognized internationally, as evidenced by the prestigious ''Scientist Medal of the Year 2017'' awarded by IAAM in Stockholm, Sweden. Moreover, he received accolades from Iranian Ministry of Science, Technology, and Research, as an ''Outstanding Emerging Researcher'' in 2017. His exceptional Ph.D. dissertation was honored as the ''Dissertation of the Year 2016'' by Iranian Academic Centre for Education, Culture, and Research in the Engineering Group. He also serves as the Secretary for the IEEE Consumer Technology Society on Machine Learning, Deep Learning, and AI. Furthermore, he assumes the role of the Track Chair of Machine Learning, Deep Learning, and AI in the CE (MDA) Track for the upcoming 2024 IEEE International Conference on Consumer Electronics (ICCE). In 2022, he chaired the ''5G and Beyond Communications'' Session at the prestigious IEEE International Conference on Communications (ICC). His expertise as a reviewer extends to leading journals and he actively participates in various conferences.

ALLAHYAR MONTAZERI (Member, IEEE) received the B.S. degree in electrical engineering from the University of Tehran, Iran, in 2000, and the M.Sc. and Ph.D. degrees in electrical engineering from Iran University of Science and Technology, Tehran, in 2002 and 2009, respectively. From 2010 to 2011, he was a Research Fellow with the Fraunhofer Institute, Germany, and continued his research with Fraunhofer IDMT and the Control Engineering Group, Ilmenau University, from 2011 to 2013. He also held visiting research positions with the Control Engineering Group, ETH Zürich, Switzerland, and the Chemical Engineering Group, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. Since 2013, he has been with the School of Engineering, Lancaster University, U.K., where he is currently an Associate Professor. His research has been supported by various U.K. research councils and industrial partners, including the Engineering and Physical Sciences Research Council (EPSRC) and the National Nuclear Laboratory. His research interests include control theory and digital signal processing, with a focus on adaptive signal processing and control, robust control, system identification (linear and nonlinear), estimation theory, and evolutionary computing and optimization. He received the ERCIM Fellowship in 2010 and the Humboldt Research Award in 2011. He is a fellow of the Higher Education Academy. His work has applications in active noise and vibration control systems and robotics. He currently serves as the Editor-in-Chief for the *Journal of Intelligent and Fuzzy Systems*, and an Associate Editor for *Frontiers in Robotics and AI and Science Progress*. He is a member of the Executive Board of U.K. Automatic Control Network and serves on the IFAC Technical Committees on ''Adaptive and Learning Systems'' and ''Modelling, Identification, and Signal Processing.''