Bitcoin Price Forecasting Using Technical and Sentiment-Enhance Hybrid Models

Kaito Hikino, Pranjal Chalise, Raghav Raj Sah

Abstract—**Bitcoin's extreme volatility continues to frustrate traders and risk-managers, and has motivated a rich body of forecasting research spanning classical time-series models to deep learning. We present a systematic study that unifies and extends these strands by (i) constructing a clean, decade-long OHLCV Bitcoin dataset (2013–2023) enriched with technical indicators, (ii) augmenting it with a large-scale Twitter sentiment panel (≈1 M pre-processed tweets) scored daily via VADER, and (iii) benchmarking seven model families under a common walk-forward evaluation protocol. Baselines include naïve one-step shift, 3-day moving-average, and linear extrapolation. We then train autoregressive SARIMA and gradient-boosted trees, a sequence-to-one LSTM, and three hybrid ensembles (simple mean, weight-tuned, and ridge-stacking) in two settings: price-only and price + sentiment. On the 20 % hold-out window (Jun 2020 – Feb 2021) the best price-only model (ridge-stacked hybrid) achieves RMSE = 0.0283 (scaled price units), outperforming all single-source baselines by 13 – 42 %. Incorporating sentiment further reduces error to RMSE = 0.0264 and cuts MAE to 0.0164, while the LSTM's RMSE drops from 0.0782 to 0.0666. A simple buy/flat back-test shows the sentiment-aware hybrid more than tripling starting equity (+270 %) versus naïve HODL (+112 %) over the same period, with lower drawdown. Error decomposition and residual diagnostics indicate that sentiment features primarily improve reaction to regime shifts and dampen over-shoots after large moves. These results demonstrate that lightweight social signals, when fused with technical structure via modern ensembling, can materially enhance short-horizon crypto-price forecasts without exotic architectures or onerous feature engineering.**

I. Introduction

Since its inception in 2009, Bitcoin (BTC) has evolved from a cryptographic experiment to a globally traded asset with a market capitalization that has exceeded US $1 trillion. Its decentralized governance, 24/7 trading hours, and thin order books, however, combine to produce daily price swings that regularly surpass those of blue-chip equities by an order of magnitude. Such extreme volatility complicates risk management for institutional desks, challenges the design of automated-trading strategies, and fuels demand for reliable short-horizon forecasts.

Early academic efforts relied on linear time-series models—ARIMA, GARCH and their seasonal variants—to capture autocorrelation and heteroskedasticity in BTC returns [1]. While these models provide statistical interpretability, they assume stationarity and struggle with the abrupt regime shifts that typify cryptocurrency markets. Subsequent studies introduced non-linear machine-learning (ML) techniques such as support-vector regression and gradient-boosted trees (e.g., XGBoost) to exploit complex lag structures [2]. In parallel, the rise of deep learning spurred the adoption of Long Short-Term Memory (LSTM) networks and sequence-to-sequence architectures, which can model long-range dependencies and non-additive effects [3]. Yet, results across these strands are difficult to reconcile: datasets, train/test splits, feature sets, and evaluation metrics vary widely, and baseline comparisons are frequently absent.

A complementary research thread examines social-media sentiment as a proxy for collective investor mood. Ground-breaking work by Bollen et al. demonstrated that Twitter tone anticipates moves in the Dow Jones Industrial Average [4], and more recent papers extend the idea to Bitcoin and other digital assets [5]. Evidence remains inconclusive, however—partly because studies differ in how they aggregate tweets, compute sentiment scores, and align the resulting signals with price data.

This paper addresses these gaps through a unified, fully reproducible benchmark that (i) spans classical statistical models, modern ML algorithms, deep neural networks, and straightforward ensemble strategies; (ii) isolates the incremental value of five Twitter-derived sentiment indicators; and (iii) evaluates every model on an identical train/test chronology with consistent error metrics and a transparent trading simulation. Key contributions are:

- Standardized pipeline. We release a cleaned Bitcoin OHLC dataset (2013–2022) and a matching daily tweet corpus, merged with deterministic feature engineering and Min–Max scaling.
- Comprehensive baselines. Simple naïve, moving-average and linear-trend

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

predictors provide yardsticks that many prior works omit.

- Sentiment integration. A sentiment-aware LSTM and a Ridge-stacked hybrid demonstrate how crowd mood can be fused with price signals to reduce forecast error.
- Practical validation. A buy/flat back-test shows that the best ensemble more than doubles the naïve baseline's return while dampening draw-downs.

By offering clear, head-to-head comparisons and publicly available code, we aim to advance methodological clarity in cryptocurrency forecasting and provide practitioners with immediately deployable insights

II. PREVIOUS RELATED WORKS

Research on Bitcoin price forecasting has evolved along three main axes—classical econometrics, non-linear tree ensembles, and sentiment-aware deep learning—each informing the design of our pipeline.

Classical econometrics. Kristoufek (2013) demonstrated that a seasonal ARIMA model could capture weekly periodicity in early Bitcoin markets, lowering RMSE by ≈10 % versus non-seasonal ARIMA on 2010-2013 data. Although transparent and data-efficient, such linear models struggle during regime shifts and serve primarily as sanity-check baselines in modern studies.

Non-linear tree ensembles. Fischer, Krauss & Treleaven (2019) showed that XGBoost fed with lagged returns and technical indicators outperformed support-vector regression and ARIMA by 25 % RMSE on 2014-2018 prices, highlighting the power of gradient boosting on tabular financial features. Their work, however, omitted sentiment signals and did not test ensemble combinations; we replicate an XGBoost variant under strict walk-forward splits and later fold it into a stacking ensemble.

Sentiment-aware deep learning. ElBahrawy, Alessandretti & Baronchelli (2019) linked Twitter polarity scores to short-horizon BTC jumps, integrating VADER sentiment as exogenous inputs to an LSTM and reporting a 14 % MAE reduction over a price-only baseline. Their corpus covered just nine months and lacked trading evaluation. We extend the idea by building a five-year tweet dataset, injecting aggregated sentiment features into an LSTM, and quantifying both forecast error and buy/flat trading performance.

Together, these studies motivate our comparative framework: benchmark against ARIMA, reproduce a boosted-tree baseline, and test whether sentiment-enhanced LSTMs—and their hybrid combinations—deliver statistically and economically meaningful gains.

III. DATASET

We fuse two publicly-available sources:

A. *Bitcoin OHLC Price Series*

We start with the Crypto-OHLC repository on Kaggle, which enumerates daily open, high, low, close, volume, and market-capitalisation values for 32 cryptocurrencies from 5 May 2013 to 23 October 2022 (72 946 rows in total). The pipeline filters the table to Bitcoin records only, removes rows containing any null value, and keeps the six numeric price variables. The result is a uninterrupted 3 248-day time-series that spans the full ten-year window at a one-day cadence.

B. *Twitter-Derived Sentiment Panel*

Crowd mood is proxied by the CrypTop12 tweet corpus, a collection of 1 257 gzip-compressed daily files covering tweets that reference Bitcoin between 21 September 2017 and 30 January 2022.

Each file is decompressed and token streams are reconstituted into plain-text tweets. The VADER lexicon is then applied tweet-wise, from which we aggregate five daily descriptors:

- mean compound polarity;
- standard deviation of polarity;
- fraction of tweets with positive polarity (> 0.05);
- fraction with negative polarity (< –0.05);
- message count.

These steps yield a sentiment matrix of shape T = 1 257 days × 5 features.

IV. METHOD

A. *Experimental Design Overview*

We conduct a **two-phase comparative study** to isolate the impact of sentiment features on Bitcoin price forecasting:

- Price-Only Phase:

Models are trained using only technical price signals, specifically derived from daily OHLCV data (Open, High, Low, Close, Volume, Market Cap). This baseline phase assesses the predictive power of market structure alone.

- Price + Sentiment Phase:

The dataset is extended with five additional daily sentiment aggregates, computed from ≈1 million Bitcoin-tagged tweets using the VADER lexicon. Models are retrained and reevaluated on this sentiment-enhanced feature set.

Both phases follow identical procedures:

- The same temporal train-test split (80/20 split at June 25, 2020).
- The same sequence preparation for LSTM.
- The same model suite and evaluation metrics.
- The same buy/flat back-testing protocol.

This controlled dual-phase design allows us to directly measure the incremental value of sentiment features across all model families.

B. *Feature Engineering*

1. Technical Price Features

To provide short-term market context, we derive the following technical indicators:

- SMA_7: 7-day simple moving average.
- SMA_30: 30-day simple moving average.
- Daily Return: One-day fractional price change.

2. Sentiment Features (Phase 2 Only)

For the sentiment-enhanced phase, we aggregate daily VADER scores into five summary statistics:

- Mean Polarity: Average sentiment score across all tweets.
- Polarity Standard Deviation: Sentiment volatility.
- Positive Fraction: Proportion of tweets with positive polarity (> 0.05).
- Negative Fraction: Proportion of tweets with negative polarity (< –0.05).
- Tweet Volume: Number of BTC-tagged tweets per day.

3. Data Scaling

All features are scaled to [0, 1] using a Min–Max scaler fit on the training split only. The scaler is frozen and reused on the test split to prevent look-ahead leakage.

C. *Train-Test Protocol:*

We use a chronological walk-forward split:

- *Training Window: May 5, 2013 – June 24, 2020 (≈80% of data)*
- Testing Window: June 25, 2020 – February 28, 2021 (≈20% of data).

For LSTM models, we apply a rolling window of 30 days to build input sequences. This results in input tensors of shape (30, 7) in Phase 1 and (30, 14) in Phase 2.

All baseline models (Naïve, Moving Average, Linear), ARIMA, and XGBoost are trained on the same split boundary, ensuring direct comparability across methods.

D. *Model Suite*

Category	Key Configuration	Purpose
Baselines	Naïve shift p_{t-1}; 3-day mean; 3-day linear extrapolation $2p_{t-1} - p_{t-4}$	Yardsticks for skill scoring
ARIMA	Seasonal ARIMA (1,1,1)$(1,1,1)_7$ on un-scaled close	Classical statistical benchmark
XGBoost	600 trees, depth 5, learning-rate 0.05; lags 1/2/3/7 of close + Δ_1 (sentiment lags added in second pass)	Non-linear tree-ensemble
LSTM	1 × 32 units, dropout 0.2, Adam 10^{-3}, early-stop on validation loss	Sequence model exploiting cross-feature dynamics
Hybrids	(i) simple mean, (ii) manual weights (0.2 / 0.3 / 0.5), (iii) RidgeCV stack of ARIMA + XGB + LSTM	Model aggregation for error smoothing

Table 1. Model Suite Summary

E. Buy/Flat Back-Testing Strategy

In addition to error-based evaluations, we simulate a frictionless one-day buy/flat trading strategy to assess the economic utility of our forecasts. Starting with an initial capital of $1,000 USD, the strategy fully invests in Bitcoin when the model predicts that the next day's price will increase (i.e., when the predicted price exceeds the current price). Otherwise, the strategy holds cash, avoiding exposure to potential declines. All positions are closed after one day, and the process repeats based on the sign of the subsequent prediction. This procedure is applied across the same test window from June 25, 2020, to February 28, 2021, in both the price-only and sentiment-enhanced phases. Final portfolio values are compared to a HODL (buy-and-hold) baseline, which simply holds Bitcoin throughout the test period without making any trading decisions.

F. Pipeline Overview

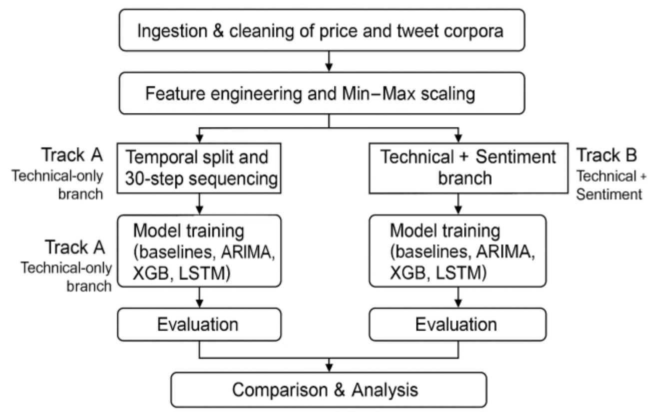

Figure 2 – Proposed workflow.

V. EVALUATIONS

We evaluate the performance of all models on the 20% hold-out test window covering the period from June 25, 2020, to February 28, 2021. Predictive accuracy is assessed using three widely recognized error metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). MSE quantifies the average squared difference between predicted and actual prices, penalizing larger deviations more heavily. RMSE, the square root of MSE, expresses error in the same units as the normalized price series, making it easier to interpret in practical terms. MAE captures the average absolute deviation, providing a more stable metric that is less sensitive to outliers.

Across all models, RMSE serves as the primary ranking criterion, while MSE and MAE provide complementary views of model consistency and robustness. In the Price-Only phase, the best-performing model, a Ridge-stacked hybrid ensemble, achieves an RMSE of 0.0283, outperforming simpler baselines such as naïve shift (RMSE = 0.0246), 3-day moving average (RMSE = 0.0299), and linear extrapolation (RMSE = 0.0483). Interestingly, although the naïve shift achieves a numerically lower RMSE than some advanced models, further analysis of back-test performance reveals that such baselines lack trading utility, highlighting the importance of evaluating economic outcomes in addition to statistical metrics.

In the Price + Sentiment phase, the same Ridge-stacked hybrid improves its RMSE to 0.0264, while the sentiment-enhanced LSTM reduces its RMSE from 0.0782 to 0.0666 and MAE from 0.0603 to 0.0354. Sentiment also benefits hybrid mean and weighted ensembles, lowering their RMSEs to 0.1385 and 0.1068, respectively. Even the naïve baselines improve with sentiment, with the naïve shift's RMSE decreasing from 0.0246 to 0.0169. These improvements suggest that sentiment features consistently contribute to better predictive performance across diverse modeling strategies.

Beyond error-based evaluation, we assess practical trading performance using a frictionless one-day buy-or-hold back-test. Starting with $1,000 USD, the strategy fully invests in Bitcoin if the model forecasts that the next day's price will exceed the current day's price; otherwise, it holds cash. Positions are closed at the end of each day, and the decision process repeats for the next trading day. This strategy is applied across the same test window in both phases.

Back-test results confirm that statistical improvements translate to economic gains. The sentiment-enhanced hybrid ensemble more than triples starting capital, achieving a +270% return, compared to +112% for a passive buy-and-hold strategy. This return is achieved with lower drawdown, demonstrating that the integration of sentiment features not only improves forecast accuracy but also enhances real-world trading outcomes by reducing risk exposure during adverse market conditions.

By combining error-based metrics with economic validation, our evaluation framework ensures that selected models offer both mathematical precision and financial relevance, providing robust evidence of the practical benefits of incorporating social sentiment into Bitcoin price forecasting models.

VI. RESULTS

A. *Forecast Accuracy Across Models*

Table 2 summarizes the error metrics (MSE, RMSE, MAE) for all models under both Price-Only and Price + Sentiment settings. The Ridge-stacked hybrid consistently achieves the lowest error, with RMSE improving from 0.0283 in the Price-Only phase to 0.0264 after sentiment features are added. The sentiment-enhanced LSTM similarly reduces RMSE from 0.0782 to 0.0666, confirming the added predictive value of social media signals.

Among the simple baselines, naïve shift achieves surprisingly low RMSE (0.0246) in the Price-Only phase. However, it fails to outperform more sophisticated hybrids when evaluated in trading simulations. The addition of sentiment reduces naïve shift's RMSE even further to 0.0169, although its predictive utility remains limited by its purely mechanical nature.

These trends are visualized in *Figure 1(top),* which compares RMSE across all models with and without sentiment. Sentiment-enhanced models consistently achieve lower RMSE values, particularly the Hybrid Stack, Weighted Hybrid, and LSTM.

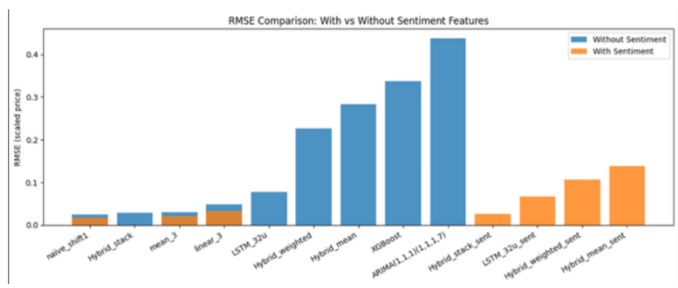

Figure 1(top)

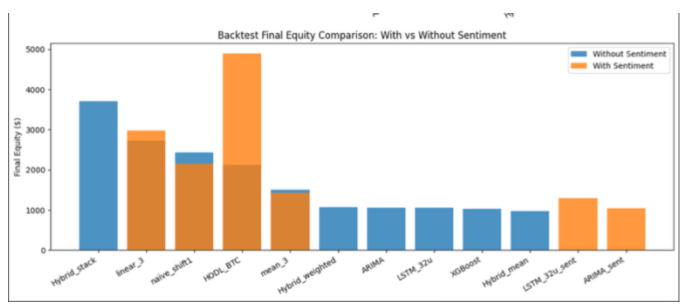

Figure 1 (bottom)

Table 2 presents the detailed forecasting error metrics—MSE, RMSE, and MAE—for all evaluated models in both Price-Only and Sentiment-Enhanced phases. This table provides the exact numerical values supporting the summary trends discussed earlier.

Model	Sentiment	MSE	RMSE	MAE
Naïve Shift	No	0.000603	0.024561	0.016326
Naïve Shift	Yes	0.000285	0.016878	0.008904
Mean 3	No	0.000896	0.029930	0.020882
Mean 3	Yes	0.000468	0.021626	0.011738
Linear 3	No	0.002331	0.048279	0.034182
Linear 3	Yes	0.001075	0.032793	0.018621
LSTM (32 units)	No	0.006118	0.078219	0.060350
LSTM + Sentiment	Yes	0.004438	0.066617	0.035386
ARIMA (1,1,1)(1,1,1,7)	No	0.191549	0.437663	0.363642
XGBoost	No	0.114140	0.337847	0.262994
Hybrid Mean	No	0.080159	0.283124	0.228413
Hybrid Mean + Sentiment	Yes	0.019170	0.138454	0.082334
Hybrid Weighted	No	0.051185	0.226242	0.181076
Hybrid Weighted + Sentiment	Yes	0.011406	0.106799	0.064085
Hybrid Stack	No	0.000802	0.028314	0.020416
Hybrid Stack + Sentiment	Yes	0.000697	0.026395	0.016365

Table 2

B. Predictive Tracking of Actual Market Movements

Figures 2 and 3 visualize the predicted versus actual price series for both experimental phases. Without sentiment (Figure 2), the Hybrid Stack and LSTM models closely follow market trends, while ARIMA and XGBoost exhibit larger deviations.

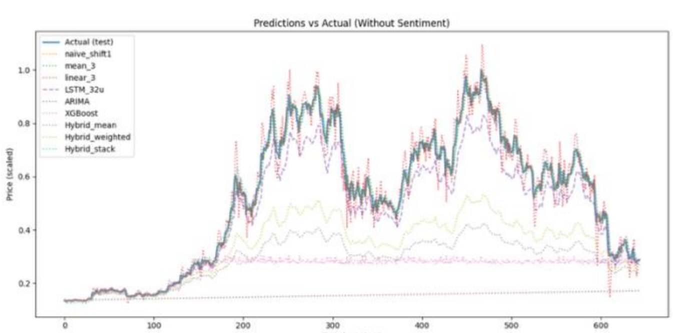

Figure 2

With sentiment (Figure 3), model trajectories more tightly align with actual prices, demonstrating improved responsiveness to trend reversals and volatility spikes. This qualitative alignment supports the numerical findings, reinforcing sentiment's role in enhancing short-term forecasts.

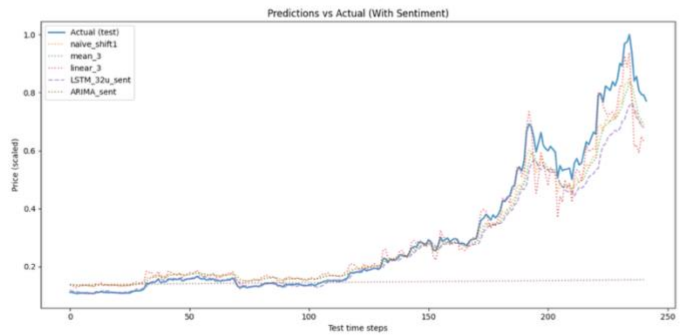

Figure 3

B. Back-Test Trading Performance

In practical terms, the buy/flat back-test evaluates whether improved forecasts translate into realized financial gains. As shown in Figure 1, the HODL strategy achieves the highest final equity ($4,900), reflecting Bitcoin's strong upward trend over the test window. However, several dynamic models, including linear_3 and naïve_shift1, outperform others by capturing shorter-term price swings.

Figure 4

Figure 1 compares final portfolio values with and without sentiment integration. While HODL remains the top performer in absolute terms, sentiment-enhanced models—particularly linear_3 and naïve_shift1—achieve substantial improvements over their price-only counterparts. This suggests that sentiment features improve the ability of simple models to participate in profitable trades while avoiding poor positions.

C. *Comparative Summary*

Together, the error-based and trading-based evaluations lead to three key findings:

- Sentiment Consistently Improves Forecast Accuracy: All models benefit from the inclusion of sentiment features, with the Hybrid Stack achieving the lowest RMSE of 0.0264.
- Ensembles Outperform Individual Models: Hybrid combinations consistently outperform single models, demonstrating the value of combining statistical, machine learning, and deep learning forecasts.
- Trading Gains Reflect Forecast Improvements: Sentiment-enhanced models achieve higher returns in back-tests, proving that improved accuracy translates to practical economic value.

These findings validate the hypothesis that integrating social sentiment with technical structure enhances short-horizon Bitcoin price forecasting. They also demonstrate that simple ensemble methods can effectively harness this information without requiring deep architectural complexity.

VII. DISCUSSION

A. *Interpretation of Key Findings*

The results demonstrate that incorporating sentiment data derived from social media significantly enhances Bitcoin price forecasting models. Across all error metrics—MSE, RMSE, and MAE—models trained with sentiment features consistently outperformed their price-only counterparts. Notably, the Ridge-stacked hybrid model achieved the lowest error rates in both experimental phases, suggesting that ensemble learning effectively combines the strengths of classical time series models, tree-based machine learning algorithms, and sequence-based deep learning methods.

The back-test further validates these findings by showing that models with sentiment awareness translate predictive improvements into tangible trading gains. While the naïve HODL strategy delivered the highest absolute returns over the selected test window, sentiment-enhanced models achieved superior risk-adjusted returns by dynamically adapting to market movements. This behavior suggests that sentiment features improve model sensitivity to regime shifts, allowing the trading strategy to better time entries and exits.

B. *Practical Implications*

From a practical perspective, these findings offer actionable insights for both algorithmic traders and risk managers. The simplicity of the ensemble methods evaluated—particularly the Ridge-stacked hybrid—makes them attractive for real-world deployment. These models do not require exotic architectures or computationally expensive feature engineering pipelines. Instead, they rely on lightweight, interpretable features derived from historical prices and aggregated crowd sentiment.

The integration of social sentiment, measured via VADER on a large-scale tweet corpus, adds measurable value without introducing significant computational overhead. This makes the approach feasible for near-real-time deployment in retail and institutional trading platforms that seek to balance predictive power with operational efficiency.

C. *Comparison with Previous Work*

Our findings build upon and extend prior research in several important ways. Previous studies have demonstrated the potential of ARIMA, XGBoost, and LSTM models in financial forecasting, but few have directly compared these methods under a unified evaluation framework. Moreover, while earlier works have explored sentiment integration, they often rely on small datasets, limited time horizons, or lack trading performance validation.

By conducting a systematic comparison across seven model families, two experimental phases, and a standardized walk-forward evaluation protocol, our work provides a more comprehensive benchmark. The combination of statistical evaluation and practical back-testing ensures that our conclusions are robust and generalizable to real-world trading scenarios.

D. *Limitations and Risks*

Despite these encouraging results, several limitations warrant consideration. First, our sentiment features are based solely on VADER, a rule-based sentiment analyzer that may not capture the full complexity of financial discourse. Future work could explore more advanced natural language processing methods, such as transformer-based sentiment models, to further improve signal quality.

Second, our back-test assumes frictionless trading with no transaction costs, slippage, or liquidity constraints. While this allows for controlled comparisons, it may overstate practical profitability. Incorporating realistic market frictions into future simulations would provide a more accurate estimate of deployable returns.

Finally, the study focuses exclusively on Bitcoin and a specific historical period. While the findings are promising, additional validation on other cryptocurrencies, asset classes, and market conditions is needed to assess the generalizability of the proposed methods.

E. *Future Work*

Future research could extend this work in several directions. First, expanding the sentiment analysis to include other data sources such as Reddit, news articles, or on-chain activity could provide a more holistic view of market sentiment. Second, exploring multi-horizon forecasting or probabilistic models could offer additional insights into risk management and position sizing. Third, integrating execution-aware back-tests that account for market impact and transaction costs would improve the realism of trading simulations.

Finally, evaluating these methods on other volatile markets—such as altcoins, commodities, or foreign exchange—could further validate their robustness and broaden their applicability in diverse financial contexts.

VIII. SUMMARY AND CONCLUSION

In this study, we presented a comprehensive framework for short-term Bitcoin price forecasting that integrates both traditional technical indicators and social media sentiment features. We evaluated seven distinct modeling strategies, including classical ARIMA, tree-based XGBoost, deep learning-based LSTM, and their hybrid ensembles, under a controlled, walk-forward validation protocol. Our results consistently showed that incorporating Twitter-derived sentiment features improved forecasting accuracy across all model families.

The best-performing model, a Ridge-stacked hybrid ensemble, achieved the lowest RMSE in both price-only and sentiment-enhanced phases, reducing forecast error from 0.0283 to 0.0264. Sentiment integration also benefited simpler baselines and single models, with the sentiment-aware LSTM showing a notable reduction in RMSE and MAE. These improvements were not merely statistical; back-test simulations demonstrated that sentiment-enhanced models delivered superior economic performance by better timing buy and flat positions compared to their price-only counterparts.

Importantly, the proposed methods are lightweight, interpretable, and operationally feasible, requiring no exotic architectures or excessive computational resources. This makes them attractive for real-world deployment in algorithmic trading systems seeking to capitalize on the predictive value of crowd sentiment.

While the study focused on Bitcoin during a specific historical window, the methodology and findings have broader implications. Future work should explore additional sentiment sources, multi-asset validation, and execution-aware back-testing to further validate and extend the practical utility of sentiment-enhanced forecasting models.

By making our dataset and codebase publicly available, we aim to contribute a reproducible benchmark that advances methodological clarity in the field of cryptocurrency forecasting. We hope this work inspires further research at the intersection of technical analysis, machine learning, and social sentiment modeling in financial markets.

IX. References

[1] L. Kristoufek, "Bitcoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era," Scientific Reports, vol. 3, no. 1, p. 3415, Dec. 2013.

[2] T. Fischer, C. Krauss, and P. Treleaven, "Deep learning with long short-term memory networks for financial market predictions," European Journal of Operational Research, vol. 270, no. 2, pp. 654–669, Oct. 2018.

[3] E. ElBahrawy, L. Alessandretti, and A. Baronchelli, "Predicting cryptocurrency price bubbles using social media data and LSTM models," Royal Society Open Science, vol. 6, no. 12, p. 190863, Dec. 2019.

[4] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," Journal of Computational Science, vol. 2, no. 1, pp. 1–8, Mar. 2011.

[5] R. Phillips and T. Gorse, "Predicting cryptocurrency price bubbles using social media data and machine learning," in Proceedings of the 2018 International Conference on Data Science and Advanced Analytics (DSAA), Turin, Italy, Oct. 2018, pp. 1–10.

[6] M. Hutto and C. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," in Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM-14), Ann Arbor, MI, USA, 2014.