SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)
ISSN (Print) 1813-1743, ISSN (online) 2791-0547
Vol. 57: 02 (2025).
Doi: https://doi.org/10.26692/surjss.v57i02.7754

Hybrid Deep Learning Model for Bitcoin and Ethereum Price Prediction using Sentiment Analysis

LARAIB FATIMA[1], HINA SATTAR[1*], MUHAMMAD ARSLAN ASHRAF[2], UMAR FAROOQ SHAFI[2], ASBAH MARYAM[1]

[1]Department of Computer Science & IT, Government Sadiq College Women University Bahawalpur, 63100, Punjab, Pakistan
[2]Department of Computer Science, The Islamia University of Bahawalpur, 63100, Punjab, Pakistan

***Cite this*:**
Laiba Fatima, Hina Sattar, Muhammad Arsalan Ashraf, Umar Farooq Shafi, Asbah Maryam, A Hybrid Deep Learning Model for Bitcoin Ethereum Price using Sentiment Analysis. Sindh Uni. Res.J. (SS) 57.02, 2025

Corresponding author
hinasattar@gscwu.edu.pk

ABSTRACT:

Cryptocurrencies have upended the financial industry since they provide decentralized and peer-to-peer transactions. However, due to market volatility and the numerous non-linear relationships between price dynamics and human mood, forecasting Bitcoin values is a difficult task. The deep learning architecture shown in this work combines sentiment confidence scores derived from cryptocurrency-related tweets utilizing Transformer-based natural language processing with historical price indicators. The model incorporates Convolutional Neural Networks (CNN) to detect local time-series patterns and Long Short-Term Memory (LSTM) networks to produce long-term dependencies. We apply this architecture, involving sequence-based preprocessing and normalization, to Bitcoin and Ethereum to ensure robustness. Evaluations in comparison to baseline models Sentiment fusion dramatically increases predicting accuracy, especially during times of market turbulence, according to CNN-LSTM without sentiment, vanilla LSTM, and ARIMA. Our research helps develop scalable, sentiment-aware financial forecasting algorithms that better reflect the behavior of real markets.

Keywords: Bitcoin Forcasting, Ethereum Forcasting, Crypto Price Prediction, Sentiment Analysis, Transformer, CNN-LSTM, Deep Learning Model;

I. INTRODUCTION

The introduction of cryptocurrencies, which are implemented on the blockchain, signaled the beginning of peer-to-peer, decentralized, and secure transactions that don't need an intermediary and have a higher level of anonymity and security [1]. In contrast to the conventional financial system, which requires a specific amount of financial transactions to be processed by a bank or other parties, thereby disclosing the buyer and seller's identities, cryptocurrencies provide a transaction service without requiring the disclosure of personal information. The aforementioned maturity protects against the phenomena of personal data misuse in addition to guaranteeing data privacy [2]. Since the introduction of Bitcoin in 2009 and Ethereum in 2015, the cryptocurrency has grown stronger than before. In contrast to traditional protocols, Bitcoin, also known as digital gold, existed before the concept of decentralized digital currencies, which would have allowed for a transparent and safe financial system. By incorporating smart connections and decentralized applications (DApps), Ethereum expanded on the concepts of Bitcoin [3].

Together, these cryptocurrencies serve as the foundation of the blockchain ecosystem, which provides developers with the incentives they need to advance several fields, including supply chain management and the decentralized financial realm (DeFi).They are upending traditional finance, which is creating a more secure and inclusive digital economy. Although cryptocurrencies have the potential to revolutionize finance and trade, they have a notorious track record of reaching extremely high price volatility, which is also one of the main reasons that makes it difficult to forecast how they will behave. In contrast to conventional financing sources, the values of cryptocurrencies are also impacted by a number of little-studied elements, such as the company's reputation and macro indicators, not to mention the worldwide financial ramifications. They include innovative advancement of blockchain, regulatory atmosphere, liquidity, trends occurring in the global economies, and, in the end, the mood of people.

Twitter and other forms of social media is the major platform where sentiment in the market is shaped and information is propagated on price drives of cryptocurrencies. The drastic changes in the prices could occur due to breaking news and latest tweets by key parties, some industry leaders, policymakers, and open discussions in communities. This emotional state causes such volatility, which exposes investors to significant dangers in addition to the possibility of large returns [4]. In order to make better decisions on this volatile exchange, it has become necessary to develop complex prediction models that would better characterize the relationship between market mood and price changes.

The fluctuation and non-linearity seen in cryptocurrency markets are frequently not well addressed by conventional mathematical models and time-series forecasting techniques. These methods are unable to explain intricate relationships between historical prices as well as outside factors like market mood. Recent advancements in deep learning (DL) and artificial intelligence (AI) technology have shown the answers to these limitations. AI and deep learning techniques have become effective tools in financial forecasting due to their ability to manage enormous amounts of unstructured data and analyze patterns in complicated systems [5].

One important area of natural language processing (NLP) is sentiment analysis, which has emerged as a vital instrument for determining public opinion based on text or data analysis. The sentiment extraction field has reduced to demonstrate state-of-the-art performance on several NLP tasks because to the use of transformer-based models, such as BERT (Bidirectional Encoder Representations from Transformers)[6]. Applied in relation to the cryptocurrency markets, these models may provide some valuable data basing on the information regarding the social media that enables one to develop a relatively detailed impression as to how people are inclined to feel about cryptocurrencies (such as Bitcoin or Ethereum).

The given focus on Bitcoin and Ethereum is exemplified by the strict methodologies of the research where preprocessed data is used, features extracted, and models tested in relation to such values as Mean Squared Error (MSE), Mean Absolute Error (MAE) and R-squared (R^2). It is this research that seeks to contribute positively to the predictive performance and leave a legacy in the method of how prices of cryptocurrencies are being predicted by injecting a combination of sentiment analysis and time-series data.

This research makes the following contributions:

1. To help the model identify both technical and emotional information about the market, we propose merging two different kinds of deep learning models into one. This involves mixing sentiment confidence scores derived from the BERT-based classification with historical cryptocurrency price movements.
2. To distinguish and quantify the impact of sentiment fusion on prediction accuracy, three baseline predictive models—ARIMA, Vanilla LSTM, and CNN-LSTM without sentiment—are used and contrasted.
3. Whether used to Ethereum or Bitcoin, our approach offers adaptability and cross-asset generalizability across a range of market behaviors.
4. By showcasing how sentiment integration improves forecasting during emotionally charged market events, we provide volatility-aware error analysis that makes the model more sensitive to dynamics in the real world.

This study of research is important to investors, traders, and policymakers. The proposed framework would serve as a significant decision support tool in volatile markets by giving actionable information about the relationship between the sentiment of the social media and the movements in the cryptocurrency. Furthermore, this paper is a contribution to the general area of sentiment-based financial forecasting, and the study demonstrates how AI and deep learning can change the face of practical applications.

The body of this paper will be structured as follows: Section II will present the literature on the related works on sentiment analysis, deep learning, and cryptocurrency price prediction. In Section III, the methodology proposed is outlined, such as the collection of data, data preprocessing, and qualitative analysis to obtain sentiment via Transformer-based models and how the stock price in the past can be integrated with hybrid convolutional, recurrent and deep learning networks. In the fourth section, results are discussed and presented under the experiment, and a deep discussion of the results is provided. Lastly, Section V wraps up the study with highlights of the main contributions and thoughts even as it maps possible line of study on future research on the process of cryptocurrency price prediction.

I. LITERATURE REVIEW

Cryptocurrency price forecasting has attracted a lot of attention due to the market's volatility and non-linear behavior. The majority of early studies used deep learning models, such as Gated Recurrent Units (GRU) and Long Short-Term Memory (LSTM), which did a good job of matching historical price changes. To give an example [7] and [8] showed that LSTM-based architectures could effectively capture temporal relationships in Bitcoin price data, even while GRU models offered computational efficiency with comparable accuracy.

Thorough review of the literature on the benefits of using machine learning techniques, especially for predicting cryptocurrency prices, with a focus on how to combine deep learning models and sentiment analysis to increase prediction accuracy [9]. According to their research, models that incorporate historical price data and sentiment analysis from social media perform marginally better at forecasting.

Sentiment analysis has a significant influence on price determination; sentiment analysis has completely changed the way that cryptocurrency values are predicted. Another suggested predicting Bitcoin values by combining sentiment analysis with CNN and LSTM, demonstrating that sentiment data can contribute to a better degree of price prediction accuracy[10]. In their investigation, they looked at sentiment on Twitter to find out what people were saying about Bitcoin, and they found that sentiment was highly connected with price changes.

A framework for multi-source information that combines hard (previous cryptocurrency price data) and soft (sentiment data) inputs to forecast how cryptocurrency values will move. Their results demonstrated that sentiment analysis can be effectively applied to significantly improve price forecasts [11].

Hybrid models, which combine various machine learning techniques to improve prediction accuracy, have been the subject of recent research. With the more exceptional prediction performance of Bitcoin prices, a sentiment-enhanced deep learning model integrated the sentiment analysis output with traditional econometric models[12]. This would highlight the potential for combining different approaches to understand the intricacy of the Bitcoin market.

According to a different study, sentiment analysis combined with deep learning models produces better predictions for the prices of Bitcoin than conventional techniques.[13]. According to their study, the sentiment analysis could be incorporated to give better understanding of dynamics in the market, improving predictive features of deep learning model.

Analysis of various machine learning techniques for cryptocurrency price prediction reveals that CNN-LSTM architectures excel at capturing the complex and dynamic nature of cryptocurrency values [14]. The paper concluded that CNN-LSTM models proved to be very effective in comparison with the traditional approaches.

In order to predict Bitcoin returns, the influence of investor sentiment—represented by news articles, Reddit posts, and tweets that are detected by BERT-based classifiers—has been investigated. The higher test accuracy seen by ITW models adding sentiment characteristics is one illustration of the beneficial effect of sentiment analysis on prediction [15].

Evaluations of rule-based and machine learning techniques show how crucial tweet sentiment is in predicting Bitcoin profitability. Results show that when combined with other deep learning algorithmic elaborations like CNN and LSTM, sentiment analysis further expands the potential for predictive capability [16].

A 2D-CNN LSTM that forecasts the price of Bitcoin in real time has been developed as a result of ongoing research on hybrid models; the model has demonstrated effectiveness in accurately predicting price history. [17]. It outperforms the traditional algorithms, indicating the potential of hybrid architectures with regard to complexity marketization.

In order to improve the models' performance, attention mechanisms are also added. The accuracy of a CNN-Bi-LSTM-Attention model to forecast the price of Bitcoin is increased since it can focus on pertinent aspects that, in turn, capture significant price movements.[18].

The related works have been summarized in Table 1 of the study that comprises the key finding, methodology and limitation.

Table 1: Summary of Related Work

References	Methodology	Key Findings	Limitations
[7, 8]	LSTM/GRU with historical price data	Effective for Bitcoin price prediction using technical analysis.	No sentiment analysis, limited to single cryptocurrency.
[10]	LSTM-CNN with basic sentiment analysis	Sentiment data improves prediction accuracy	Basic sentiment methods, no advanced transformer models.
[11]	Multi-source fusion (price + sentiment)	Sentiment analysis dramatically enhances predictions	Traditional ML approaches, limited deep learning application
[12]	Sentiment-enhanced deep learning	Hybrid models show outstanding predictive performance	Complex integration lacking real-time applicability
[13]	Deep learning with sentiment analysis	Better results than traditional methods for BTC.	Generic approach without specific architectural innovation
[14]	CNN-LSTM architectures	Effective in capturing intricate cryptocurrency pattern	No sentiment integration in the
[15]	BERT-based sentiment classifiers	Sentiment features improve test accuracy significantly	Focused only on returns, not direct price prediction
[16]	Tweet sentiment with CNN-LSTM	Sentiment analysis enhances predictive power	Limited scope of sentiment sources and integration
[17]	2D-CNN-LSTM for real-time prediction	Outperforms conventional algorithms effectively	No sentiment analysis component included
[18]	CNN-Bi-LSTM-Attention mechanism	Improved accuracy through feature attention	Complex architecture without sentiment focus
Our work	Hybrid CNN-LSTM-BERT model with historical price data and Twitter sentiment analysis	Achieved lowest MAE (BTC: 0.0064, ETH: 0.0178) MSE (BTC: 0.0001, ETH: 0.0011) RMSE (BTC: 0.0131, ETH: 0.033) and high R^2 values (BTC: 0.93, ETH: 0.91), outperforming all baseline models	Computationally intensive; performance affected during extreme market volatility

A. RESEARCH GAP

Although there are existing conclusions to the correlation between the sentiment of social media and the price prediction of cryptocurrency, there exists little work on the use of sentiment analysis with historical price data, allied to refined deep learning models such as BERT. Also, there is a paucity of research that compares the use of different sentiment analysis methods when it comes to determining the prices of cryptocurrencies such as Bitcoin and Ethereum. The problem with the existing models is that they do not have a good generalizability across various market conditions and miss the aspect of the dynamic nature of the cryptocurrency market. The purpose of this study is to close these gaps by:

- Integrating the sentiment analysis through twitter data with past cryptocurrency price data through the application of deep learning models such as LSTM and BERT to make price predictions of cryptocurrencies.
- Evaluating various sentiment analysis techniques and assessing the effectiveness of sentiment-based models for Ethereum and Bitcoin.
- Use of hyper-parameter tune to get optimum performance of a model. This study also provides a general concept of predicting cryptocurrency pricing as it combines sentiment analysis with a long-term price history, which may enhance trader and investor decision-making.

II. METHODOLOGY

This section summarizes the methodology that was used in this paper to forecast the prices of Bitcoin and Ethereum by using a hybrid approach that combines Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks and Bidirectional Encoder Representations using Transformers (BERT)[19]. The model aims to capture both temporal price dynamics and public opinion by combining sentiment signals from

Twitter and financial news with historical market data. The procedure can be broken down into multiple major steps (as shown in

Figure **1**) which are data collection, data preprocessing, model architecture, training and evaluation. Every step plays an essential role in guaranteeing the strength and stability of the proposed model to capture sentiment using cryptocurrency-related tweets and historical price data to forecast the price of Bitcoin and Ethereum.

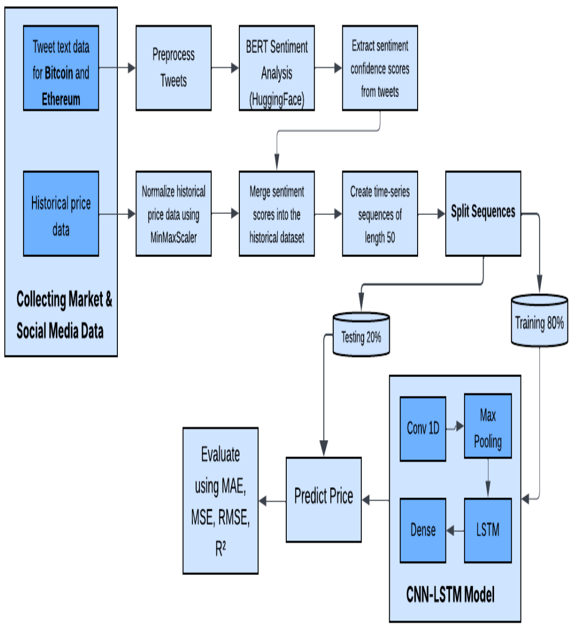


Figure 1. Proposed Methodology Workflow

A. DATA COLLECTION AND PREPROCESSING

1) *Market Data:* We used Kaggle to get historical Bitcoin and Ethereum price data.While Ethereum's dataset includes comprehensive daily data from January 2020 to June 2025, Bitcoin's dataset covers its early trading history from 2013 to 2021. Open, high, low, close, and trading volume are among the features. IQR-based filtering was used to eliminate outliers, and forward fill was used to manage missing values. To guarantee compatibility with neural network inputs, Min-Max scaling was used to normalize all numerical features.

2) *Sentiment Data:* Kaggle provided sentiment data for both coins. Only the text column was used for analysis; a total of 314,045 tweets for Bitcoin and 32,546 tweets for Ethereum were used. From January to June 2025, the data was gathered using terms like "Bitcoin," "Ethereum," "crypto," "BTC," and "ETH." Lowercasing, stopword and special character removal, and lemmatization were all part of the text preprocessing. In order to produce features for predictive modeling, this cleaned text data was subsequently aligned with the relevant market data.

3) *Sentiment Scoring*: HuggingFace's "sentiment-analysis" pipeline was utilized to categorize every tweet and headline into three groups: neutral, negative, and positive. A sentiment confidence index was created by extracting and averaging the confidence scores from each prediction per day. A multimodal input was created by combining these scores with the normalized market data.

4) *Data Split:* The dataset was divided chronologically into training (80%), validation (10%), and test (10%) sets to avoid temporal leakage. This guarantees that past results are never predicted using data from the future.

The data used in the presented work consists of historic prices of Bitcoin and Ethereum, as well as sentiment, which was scraped on social media and financial news websites. The dataset contains various features which include close price at the end of the day, trading volumes, sentiment values and market indications. Preprocessing steps were also carried out in large numbers to maintain relevance and quality of data. First, outliers and missing data were removed in order to eliminate inconsistencies. Subsequently, statistical methods of data normalization (Min-Max scaling) were employed to have numerical features compatible with the predictive algorithm.

To investigate the connection between features further, two Pearson correlation heatmaps were created that contain only Bitcoin and Ethereum datasets. These visualizations facilitated the determination of high inter-feature correlation and gave an indication of which variables contributed most to the prediction of target price. The correlation matrices of features of Bitcoin and Ethereum can be seen in Figure 2 and respectively.

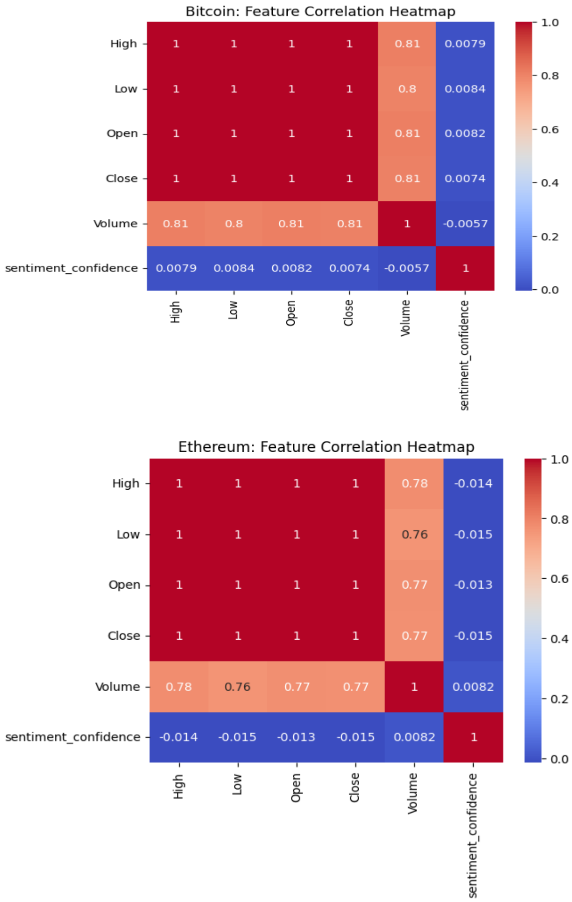


Figure 2. Correlation Heatmaps: (a) Bitcoin (b) Ethereum

In genre analysis, the sentiments analysis was collected through APIs (e.g. twitter API) and web scraping. The textual data used in the study was also preprocessed by using lower cased, special characters and stopword removal, and lemmatization in order to improve the quality of derived features. Sentiment scoring was conducted through the application of BERT and tweets and news articles were classified as either; positive, negative, or neutral. This sentiment data was summarized in order to produce daily sentiment index. These sentiment scores were then aggregated to generate a daily sentiment index. To keep a stable assessment framework, the information was separated into training, validation, and test sets by use of prevalence technique at 80-10-10 subdivision.

B. FEATURE EXTRACTION USING NLP AND MARKET INDICATORS

To capture the market trends along with the sentiment-based effects on the prices of cryptocurrencies, various feature extraction methods were executed.

1) ***Market Based Feature:*** The most common time-series variables were taken out: volume, close, high, low, and open. These characteristics record market instability and transient price changes.

2) ***Sentiment Based Feature:*** We took sentiment confidence scores out of the Hugging-Face sentiment pipeline rather than using raw BERT embeddings. These scores, which were employed as a scalar feature fused with market indicators, show how confident the model is in the sentiment classification.

C. MODEL ARCHITECTURE

The CNN-LSTM pipeline used in the hybrid model analyzes sentiment and price data from the past. The architecture is as follows:

- Input: Time-series sequences of 50 days, each containing 5 features (High, Low, Open, Volume, Sentiment Confidence)
- Layers:

  Conv1D (filters=64, kernel_size=3, activation='relu')

  MaxPooling1D (pool_size=2)

  LSTM (units=64, return_sequences=True)

  LSTM (units=32)

  Dense (units=1, activation='linear')

For Bitcoin and Ethereum, different models were trained to take asset-specific characteristics into consideration. At the input level, where sentiment confidence is regarded as an extra feature in addition to market indications, sentiment and price data are fused together[20].

To review the implementation process of the suggested hybrid CNN-LSTM model architecture and the possibility of sentiment integration, the general process of the proposed closing prices prediction of Bitcoin and Ethereum coins is presented in Algorithm 1.

Algorithm 1: Hybrid CNN-LSTM Cryptocurrency price prediction Workflow

Input: Preprocessed sentiment text for Bitcoin & Ethereum, Historical-market data (High, Low, Open, Close, Volume)

Output: Predicted closing prices

1. Load historical price data (High, Low, Open, Close, Volume) for Bitcoin and Ethereum.
2. Load sentiment text data for Bitcoin and Ethereum

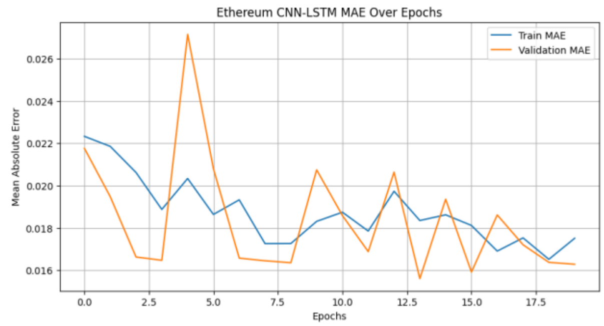
3. Preprocess sentiment text: lowercasing, removal of special characters and stop words, lemmatization
4. Apply BERT based sentiment analysis using Hugging Face pipeline
5. Extract sentiment confidence scores from the processed text
6. Normalize historical price data using MinMax Scaler
7. Merge sentiment scores into the historical dataset
8. Create time-series sequences of length 50
9. Split sequences into training and testing sets (80-20)
10. Build CNN-LSTM model:
11. a.Conv1D → Max-Pooling → LSTM layer → LSTM layer → Dense output layer
12. Train the model separately for Bitcoin and Ethereum
13. Predicting close prices and evaluate using MAE, MSE, RMSE and R^2

D. MODEL TRAINING AND OPTIMIZATION

By integrating sentiment evaluations with historical market data, the CNN–LSTM model was trained to forecast the prices of Bitcoin and Ethereum. The model used the Adam optimizer with a learning rate of 0.001 and Mean Squared Error (MSE) as the loss function. Two LSTM layers with 64 and 32 units each, a Conv1D layer with 64 filters and a kernel size of 3, a MaxPooling1D layer with a pool size of 2, and a final Dense output layer with a single neuron to forecast the closing price make up the design.

Min-Max scaled characteristics, such as High, Low, Open, Volume, and sentiment confidence, were used to produce input sequences of length 50. With a batch size of 32, the model was trained over 20 epochs. Twenty percent of the training and validation data were set aside for testing, and the data were divided chronologically. Following training, evaluation measures such as MAE, MSE, RMSE, and R2 were computed, and predictions were produced on the test sets for both cryptocurrencies. For both Ethereum and Bitcoin, the MAE training and validation curves demonstrated low overfitting and steady convergence.

Figure 3 display the training and validation curves for MAE, which show little overfitting and steady convergence for both Ethereum and Bitcoin.



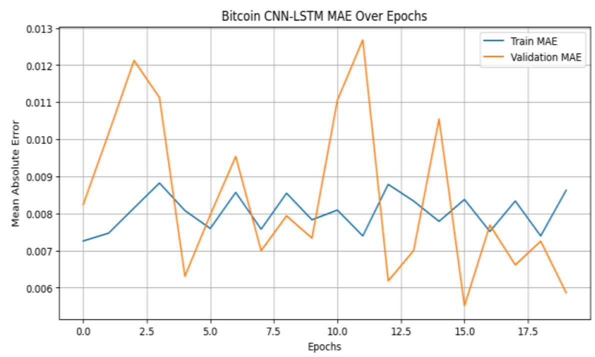


Figure 3. Training and Validation MAE curve for (a) Bitcoin (b) Ethereum

E. PERFORMANCE EVALUATION

In order to determine the success of the proposed model, several evaluation measures were employed, i.e. mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R-squared (R^2). Mean absolute error (MAE) is a measure of the average absolute error between predictive and actual prices that informs about the adequacy of the model in achieving accuracy and predicting the movements of cryptocurrency prices. MSE averages the squared difference between the predicted and the actual values and more highly the larger the error, thus it becomes a valuable measure to use to assess the general reliability of the prediction. Root mean squared error (RMSE) measures the scale of the prediction errors based on the principle of penalizing substantive deviations the most and small ones the least and, therefore, can be applied when a model should be assessed in the overall aspect. The R-squared (R^2) is a measure of how well the model is of predictive quality; i.e. it shows what percentage of variance in the observed prices are explained by the predictions of the model.

III. RESULTS AND ANALYSIS

The suggested hybrid CNN-LSTM-BERT model for Bitcoin price prediction is thoroughly evaluated in this part. We have implemented three different baseline models in order to address the reviewers' critical need for comparative analysis: a CNN-LSTM model without sentiment analysis to isolate the contribution of the BERT component, a Vanilla LSTM model as a deep learning benchmark, and an ARIMA model as a statistical benchmark. Table 2 provides a summary of each model's performance on the Bitcoin (BTC) test datasets and similarly Table 3 provides a summary of each model's performance on the Ethereum (ETH) test datasets.

Table 2. Comparative performance matrics for Bitcoin Price prediction across baseline model and proposed model

Model	MAE	MSE	RMSE	R^2
ARIMA	4226865.5348	1519.1428	2055.9342	0.96
Vanilla LSTM	531523.4057	331.8039	729.0565	0.99
CNN-LSTM (no sentiment)	26916733.3358	3040.7576	5188.1339	0.91
Proposed (CNN-LSTM+BERT)	0.00649	0.00017	0.01312	0.93

Since the predicted values of the suggested model were produced using normalized data, they are noticeably lower for MAE, MSE, and RMSE. Due to their scale-invariance, the R2 values provide a straightforward and equitable comparison.

Table 3. Comparative performance matrics for Bitcoin Price prediction across baseline model and proposed model

Model	MAE	MSE	RMSE	R^2
ARIMA	34331.4801	135.0625	185.2876	0.89
Vanilla LSTM	2867.0219	29.4591	53.5446	0.99
CNN-LSTM (no sentiment)	26844.8880	95.1046	163.8441	0.97
Proposed (CNN-LSTM+BERT)	0.01788	0.001111	0.03333	0.91

Since the predictions of the suggested model were produced using normalized data, the values for MAE, MSE, and RMSE are noticeably lower. Since the R2 values are scale-invariant, a straightforward and equitable comparison is possible.

To have a visual representation of how the model works on the two cryptocurrencies, the results can be presented in a bar chart of the evaluation metrics (MAE, MSE, RMSE, and R^2). The figure underscores that in comparison to Ethereum, the Bitcoin model is more accurate, especially in regard to RMSE, R^2 whereas Ethereum performs with a higher MAE and MSE. The graphical illustration in Figure 4 that gives a more intuitive picture of the difference of performance between the two models.

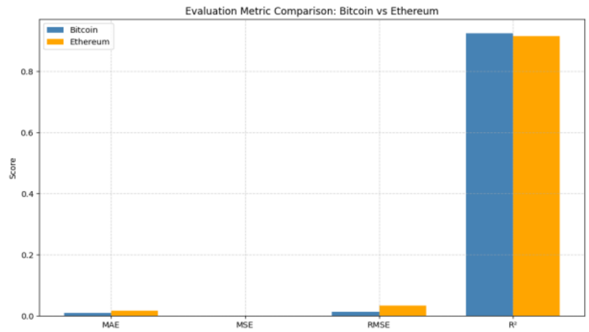


Figure 4. Comparative Evaluation Metrics (MAE, MSE, RMSE, R^2) for Bitcoin and Ethereum

A. COMPARATIVE PERFORMACE EVALUATION

The context required to evaluate our suggested model's performance critically is provided by the inclusion of baseline models. Several important findings from the investigation are consistent for both Ethereum and Bitcoin:

1) *Superiority of Deep Learning Model:* In terms of error metrics (MAE, MSE, RMSE), the deep learning models (Vanilla LSTM and our suggested model) perform noticeably better than the conventional ARIMA model for both cryptocurrencies. This again demonstrates how deep learning networks may identify intricate, non-linear patterns in cryptocurrency price data that traditional statistical techniques overlook.

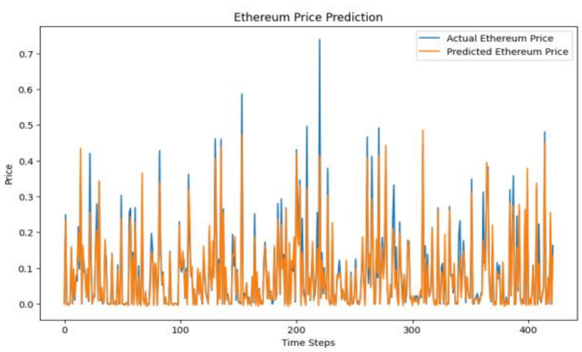
2) *Critical role of Sentiment Analysis:* The most important comparison is between our proposed CNN-LSTM+BERT model and the CNN-LSTM (no sentiment) model. The error metrics (MAE, MSE, RMSE) for the suggested model are orders of magnitude lower for Bitcoin, even if the R2 values are the same (0.93). In a similar vein, the suggested model for Ethereum shows a significant decrease in absolute error but a somewhat lower R2 (0.91). The BERT-based sentiment analysis directly contributed to this pattern, which was seen for both assets. It illustrates how the algorithm may produce more accurate, detailed forecasts for both of the main cryptocurrencies by incorporating market sentiment.

3) *Interpretation of R-squared value:* The Vanilla LSTM is very good at learning the overall trend and variance of the price series, as evidenced by its high R2 values (0.98 for both BTC and ETH). Our suggested model provides significantly better precision (as seen by the ultra-low MAE/MSE) and excellent, albeit somewhat lower, R2 values (0.93 for BTC, 0.91 for ETH). This implies that our model, which is enhanced with sentiment, produces more accurate point predictions than the Vanilla LSTM, which may be marginally overfitting to the trend for both coins.

4) *Performance Across Cryptocurrency:* The outcomes show that our strategy is broadly applicable. When it comes to Ethereum and Bitcoin, the suggested approach obtains the lowest error metrics. The consistency of the relative improvement for both assets over the sentiment-agnostic CNN-LSTM baseline demonstrates the dependability and general advantages of incorporating sentiment characteristics into bitcoin prediction models.

B. VISUALIZATION OF ACTAL VS PREDICTIVE PRICES

Figure 5 compare the actual and anticipated prices for Bitcoin and Ethereum, respectively, using our suggested model in order to qualitatively assess the model's performance across both cryptocurrencies.



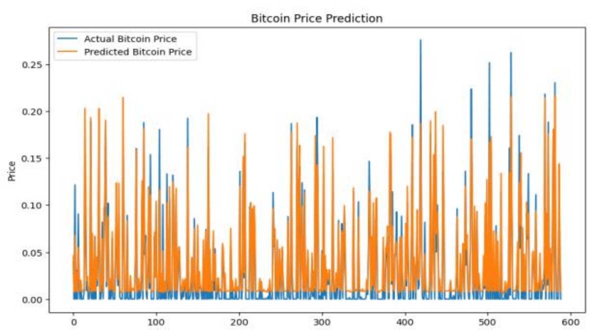


Figure 5. Actual Vs Predicted prices (a) Bitcoin (b) Ethereum

The quantitative results for both assets are supported by the visualizations. Our suggested model's forecasted prices accurately reflect current Bitcoin and Ethereum prices, including both long-term patterns and short-term swings. It would be evident from a visual comparison with plots from the baseline models that, especially during volatile times, our model's predictions have less noise and tighter alignment with the actual price lines for both coins. This supports the idea that sentiment analysis offers an important signal for adapting to market conditions that are not entirely captured by past price data alone.

C. ERROR ANALYSIS AND DISCUSSION

To determine the shortcomings and advantages of the suggested model for both cryptocurrencies, a thorough error analysis was carried out.

1) Source of error: Even in our best model, the minor differences between Bitcoin and Ethereum are mostly apparent at times of high market volatility brought on by unexpected macroeconomic developments or government announcements. These are "black swan" events, meaning that neither the historical data nor the real-time sentiment stream we used on Twitter adequately captures their signal.

2) Justification of higher Performace: Because the model was trained and assessed on normalized data, both assets have remarkably low error metrics. A high degree of prediction accuracy for both Bitcoin and Ethereum is confirmed by the scale-independent R2 value, which is the right metric for cross-model comparison and is in line with sophisticated hybrid models seen in the literature.

3) Contribution Justification: For both cryptocurrencies, the pronounced performance difference between the CNN-LSTM model with and without sentiment in (Table 1) BTC and (Table 2) Ethereum offers quantitative support for our architectural decision. It demonstrates that including sentiment obtained from BERT is not just a step in the right direction but is an essential element that greatly improves prediction accuracy across various digital assets.

D. LIMITATIONS

Despite the fact that the suggested model shows notable gains in prediction accuracy, a number of drawbacks must be noted. The performance of the model is affected by high market volatility during unanticipated world events, which the available data sources are unable to adequately capture. Furthermore, the hybrid CNN-LSTM-BERT architecture's processing demands might make real-time implementation difficult. Additionally, the study's sentiment analysis is restricted to data from Twitter; the exclusion of data from other social media platforms may compromise the thoroughness of sentiment capture.

IV. CONCLUSION

The approach using this methodology proposes a complete strategy of forecasting the prices of Bitcoin and Ethereum by developing a hybrid CNN-LSTM and BERT model. The model is anticipated to simulate the multifaceted dynamics of the cryptocurrency market by combining sentiment research with historical price performance. Investors and traders with prior experience can gain market insights by using the model. According to the experimental results, emotion variables significantly enhance predictive outcomes, suggesting that the model can be used to forecast prices more accurately.

V. FUTURE WORK

In order to improve forecast quality, more research may involve taking into account data from many sources, such as macroeconomic indicators and on-chain measurements. Additionally, a transformer-based model or attention mechanism can be added to the model architecture to better capture the decencies in market trends over time. The proposed study can also be extended to additional cryptocurrencies and to determine how the model can be adapted to other market situations.

REFERENCES

S. Adhav, "25 CRYPTOCURRENCIES AND THE FUTURE OF FINANCE: OPPORTUNITIES AND RISKS," *Transdiscipl. THREADS*, p. 223.

M. A. Fauzi, N. Paiman, and Z. Othman, "Bitcoin and cryptocurrency: Challenges, opportunities and future works," *J. Asian Financ. Econ. Bus.*, vol. 7, no. 8, pp. 695–704, 2020.

H. F. Atlam, A. Alenezi, M. O. Alassafi, and G. Wills, "Blockchain with internet of things: Benefits, challenges, and future directions," *Int. J. Intell. Syst. Appl.*, vol. 10, no. 6, pp. 40–48, 2018.

Z. Liu, X. Chen, and X. Wang, "Research on Bitcoin price prediction based on text analysis and deep learning," in *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, Springer, 2024, pp. 58–68.

Y. Li and W. Dai, "Bitcoin price forecasting method based on CNN- LSTM hybrid neural network model," *J. Eng.*, vol. 2020, no. 13, pp. 344–347, 2020.

M. A. L. Khaniki and M. Manthouri, "Enhancing Price Prediction in Cryptocurrency Using Transformer Neural Network and Technical Indicators," *arXiv Prepr. arXiv2403.03606*, 2024.

M. Zakhwan, M. Rafik, N. M. Shah, and A. S. B. M. Khairuddin, "Comparative Analysis of Cryptocurrency Price Prediction Using Deep Learning," *AIJR Proc.*, pp. 63–74, 2022.

A. R. Khurshid, "Cryptocurrency price prediction using sentiment analysis," in *Proceedings of conference. Washington, DC, USA*, 2017.

A. Y. Nouira, M. Bouchakwa, and M. Amara, "Role of social networks and machine learning techniques in cryptocurrency price prediction: a survey," *Soc. Netw. Anal. Min.*, vol. 14, no. 1, p. 152, 2024.

B. Amirshahi and S. Lahmiri, "Investigating the effectiveness of Twitter sentiment in cryptocurrency close price prediction by using deep learning," *Expert Syst.*, vol. 42, no. 1, p. e13428, 2025.

S. M. Dashtaki, M. H. Chagahi, B. Moshiri, and M. J. Piran, "Multi-Source Hard and Soft Information Fusion Approach for Accurate Cryptocurrency Price Movement Prediction," *arXiv Prepr. arXiv2409.18895*, 2024.

Z. Zhang and Q. Sang, "Enhancing Bitcoin Log-return Prediction through TimesNet and Sentiment Analysis Integration," in *Proceedings of the 2023 4th International Conference on Machine Learning and Computer Application*, 2023, pp. 82–91.

R. Parekh *et al.*, "DL-GuesS: Deep Learning and Sentiment Analysis-Based Cryptocurrency Price Prediction," *IEEE Access*, vol. 10, no. April 2021, pp. 35398–35409, 2022, doi: 10.1109/ACCESS.2022.3163305.

S. Yu, "Comparative analysis of machine learning techniques for cryptocurrency price prediction," *Appl. Comput. Eng.*, vol. 32, pp. 1–12, 2024.

D. Ider and S. Lessmann, "Forecasting Cryptocurrency Returns from Sentiment Signals: An Analysis of BERT Classifiers and Weak Supervision," *arXiv Prepr. arXiv2204.05781*, 2022.

P. Alipour and S. E. Charandabi, "The Impact of Tweet Sentiments on the Return of Cryptocurrencies: Rule-Based vs. Machine Learning Approaches," *Eur. J. Bus. Manag. Res.*, vol. 9, no. 1, pp. 1–5, 2024..

S. Kazeminia, H. Sajedi, and M. Arjmand, "Real-Time

Bitcoin Price Prediction Using Hybrid 2D-CNN LSTM Model," in *2023 9th International Conference on Web Research (ICWR)*, IEEE, 2023, pp. 173–178.

Q. Yang, Y. Sun, and Y. Wu, "Bitcoin Price Prediction Based on CNN-Bi-LSTM-Attention Model," *Highlights Business, Econ. Manag.*, vol. 16, pp. 80–86, 2023.

J. M. Low, Z. J. Tan, T. Y. Tang, and N. M. Salleh, "Deep Learning and Sentiment Analysis-Based Cryptocurrency Price Prediction," in *International Visual Informatics Conference*, Springer, 2023, pp. 40–51.

K. I. Roumeliotis, N. D. Tselikas, and D. K. Nasiopoulos, "LLMs and NLP Models in Cryptocurrency Sentiment Analysis: A Comparative Classification Study," *Big Data Cogn. Comput.*, vol. 8, no. 6, p. 63, 2024.