Beyond Trading Data: The Hidden Influence of Public Awareness and Interest on Cryptocurrency Volatility

Zeyd Boukhers  
Fraunhofer Institute for Applied Information Technology  
University of Koblenz  
Germany  
zeyd.boukhers@fit.fraunhofer.de

Cong Yang  
Soochow University  
Suzhou, China  
cong.yang@suda.edu.cn

Azeddine Bouabdellah  
University of Koblenz  
Koblenz, Germany  
bouaballahzeddine@gmail.com

Jan Jürjens  
University of Koblenz  
Fraunhofer Institute for Software and Systems Engineering  
Germany  
juerjens@uni-koblenz.de

ABSTRACT

Since Bitcoin first appeared on the scene in 2009, cryptocurrencies have become a worldwide phenomenon as important decentralized financial assets. Their decentralized nature, however, leads to notable volatility against traditional fiat currencies, making the task of accurately forecasting the crypto- fiat exchange rate complex. This study examines the various independent factors that affect the volatility of the Bitcoin-Dollar exchange rate. To this end, we propose CoMForE, a multimodal AdaBoost-LSTM ensemble model, which not only utilizes historical trading data but also incorporates public sentiments from related tweets, public interest demonstrated by search volumes, and blockchain hash-rate data. Our developed model goes a step further by predicting fluctuations in the overall cryptocurrency value distribution, thus increasing its value for investment decision-making. We have subjected this method to extensive testing via comprehensive experiments, thereby validating the importance of multimodal combination over exclusive reliance on trading data. Further experiments show that our method significantly surpasses existing forecasting tools and methodologies, demonstrating a 19.29% improvement. This result underscores the influence of external independent factors on cryptocurrency volatility.

CCS CONCEPTS

• Computing methodologies → Ensemble methods; Supervised learning by classification; • Applied computing → Economics; Forecasting; • Human-centered computing → Social media.

KEYWORDS

Cryptocurrency Forecasting, Public Awareness, Trading Data, Ensemble Learning, Deep Neural Networks

1 INTRODUCTION

Since the mining of the first Bitcoin’s genesis block in 2009, cryptocurrencies have redefined the landscape of financial assets. By January 2022, the total market capitalization of major cryptocurrencies nearly reached an astounding $1 trillion. The global cryptocurrency market exhibits a Compound Annual Growth Rate (CAGR) of 30% from 2019 to 2026. This remarkable growth not only offers lucrative investment and trading opportunities but also contributes to further expansion of the cryptocurrency market, fostering a virtuous cycle. In contrast to traditional FIAT currencies, which are regulated by central banks and governing bodies, cryptocurrencies are entirely decentralized. Transactions are validated and processed via a cryptographic network of nodes and are logged on a blockchain – a digital transaction ledger. These unique properties make forecasting the fiat-crypto exchange rate (or simply, cryptocurrency price forecasting) an exceptionally complex task.

Recent studies focused on forecasting cryptocurrency prices have predominantly leveraged neural networks due to their exceptional performance in related tasks. As a result, these approaches have achieved superior outcomes compared to traditional machine learning and statistical methods. However, a significant limitation of existing methods is that they consider only a handful of factors influencing the cryptocurrency market. The common practice is to infer a predictive function from available training sets and then evaluate the derived functions based on their generalization capabilities, presuming that price series often display homogeneous nonstationarity.

Recent studies focused on forecasting cryptocurrency prices have predominantly leveraged neural networks due to their exceptional performance in related tasks. As a result, these approaches have achieved superior outcomes compared to traditional machine learning and statistical methods. However, a significant limitation of existing methods is that they consider only a handful of factors influencing the cryptocurrency market. The common practice is to infer a predictive function from available training sets and then evaluate the derived functions based on their generalization capabilities, presuming that price series often display homogeneous nonstationarity.

In reality, cryptocurrency volatility is influenced by a multitude of factors, distinct from those affecting foreign exchange rates. Factors such as the cryptocurrency mining process’s hash rate and public awareness play a critical role in price volatility. For example, studies have revealed a sensitivity of cryptocurrency prices to public opinion sentiments. Yang et al. proposed that social media sentiment serves as a valuable predictor of future Bitcoin price volatility. Akbıyıt et al. [3] have proved this hypothesis 1

1https://gadgets.ndtv.com/cryptocurrency/news/bitcoin-price-btc-cryptocurrency-market-crash-usd-1-trillion-coinmarketcap-2726233

2https://www.globenewswire.com/news-release/2021/04/12/2208331/0/en/At-30-CAGR-Cryptocurrency-Market-Cap-Size-Value-Surges-to-Record-5-190-62-Million-by-2026-Says-Facts-Factors.html
in their study. Therefore, a holistic approach that considers these wider influences is essential for understanding the volatility of cryptocurrencies.

To enhance forecasting accuracy, this paper proposes a comprehensive utilization of all available factors that influence cryptocurrency prices. Specifically, we introduce CoMForE: Comprehensive Multimodal Crypto Forecasting Ensembler, an ensemble of multimodal Long Short-Term Memory (LSTM) models for cryptocurrency volatility prediction. This method employs trading data, social media sentiment analysis, blockchain data (including hash rate and network difficulty), and search volumes from search engines. Our focus is on Bitcoin due to its dominant market presence, its popularity among the 7812 existing cryptocurrencies, and the ample data available for analysis. Ultimately, our goal is to facilitate sound investment decision-making through the provision of reliable price forecasts and volatility distribution assessments. The following research questions guide our exploration:

- **RQ1**: Which combinations of data modalities exert the most significant influence on cryptocurrency volatility?
- **RQ2**: Does the application of ensemble learning to multimodal data yield effective outcomes in the forecasting of cryptocurrency prices?
- **RQ3**: How can we effectively interpret cryptocurrency price forecasting in conjunction with volatility distribution?

To this end, the main contributions of this paper can be summarized as follows:

- Our research encompasses factors influencing cryptocurrency volatility, including trading data, social media sentiments, blockchain metrics, and search engine volumes.
- We introduce an LSTM-based multimodal ensemble learning model that outperforms existing models, delivering reliable decision-making support to investors.
- Beyond price prediction, our model provides fluctuation distribution, enhancing investors’ understanding of forecast certainty.
- Comprehensive experiments and analyses affirm the robustness of our approach.
- We provide an open-source implementation of our model for further development and improvement.

2 RELATED WORK

In this section, we review the related works divided into three categories:

2.1 Traditional Market Price Forecasting

The task of forecasting stock prices has long been a subject of interest for scholars in the realms of finance, statistics [12], and data science [22, 39]. The aim is to empower investors with the tools necessary to augment their profits while mitigating potential losses. One of the early approaches adopted was the application of statistical models. An in-depth comparative study was conducted by Haviluddin et al. [5], examining statistical and machine learning techniques for short-term forecasting using time-series data. The study primarily compared the statistical method ARIMA, Neural Networks, and genetic algorithms. The findings underscored the superior efficiency and reliability of Neural Networks for short-term time series forecasting.

Over the past decade, the advancements in deep neural networks across diverse applications have brought them into the spotlight for financial and economic forecasting, including time-series predictions [13, 17, 29, 32]. Deep learning approaches have demonstrated their ability to significantly outperform other methods, largely due to their capability to learn concealed features of time series and historical market trends. Traditional markets are governed by relatively well-understood factors and rules that directly influence price trends, such as the number of asks, bids, and transactions. This relative clarity facilitates deep learning models in pattern recognition from historical data, leading to highly reliable forecasting, as evidenced by the high accuracy and low error rates observed when applying price, ask, and bid time series alone [29, 32]. However, attempts to apply these cutting-edge deep learning methods—proven effective in traditional markets—to the task of forecasting cryptocurrency prices have not yielded satisfactory results [24]. The distinct characteristics of the cryptocurrency market render it a more complex domain for forecasting.

2.2 Machine Learning for Cryptocurrency Price Forecasting

Cryptocurrency market volatility has driven the development of tailored forecasting approaches. Yiying and Yaze [38] proposed LSTM and ANN architectures using price, ask, and bid time series for short and long-term price predictions of Bitcoin, Ethereum, and Ripple. The results showed the LSTM’s effectiveness in capturing short-term dynamics.

McNally et al. [25] evaluated LSTM and Bayesian-optimized Recurrent Neural Network models for Bitcoin price forecasting, concluding that the LSTM had marginally superior accuracy. However, they noted the challenge of balancing overfitting and underfitting due to the high volatility of cryptocurrency time series.

Kumar and Rath [20] used only historical trading data for Ethereum price predictions, indicating LSTM’s slight outperformance over MLP for short-term forecasts. Meanwhile, Pintelas et al. [30] found LSTM-based and CNN-based models demonstrating almost random walk processes, suggesting the exploration of new approaches.

Chevalier et al. [8] introduced an AdaBoost-based approach for cryptocurrency forecasting that significantly improved performance. Despite its simplicity, AdaBoost outperformed other models like ANNs, LSTMs, KNNS, and SVMs, highlighting its generalizability and interpretability.

2.3 Sentiment Analysis and Multimodality for Cryptocurrency Price Forecasting

Krisoufek [19] found that the cryptocurrency market is influenced by various factors, including the number of asks, bids, exchange rates, and notably, public awareness, as confirmed by Akbiyik et al.[3]. Recognizing the challenge of quantifying public awareness, individual sentiments can serve as proxies, which are believed to affect price trends[31]. Young et al. [18] proposed that cryptocurrency forum sentiments might influence Bitcoin prices. They proposed a model using exclusively sentiment data and found a correlation between price trends and forum sentiments. Leveraging public opinion for finance is not a new concept, and it extends beyond
cryptocurrency. Several approaches have been suggested to utilize public sentiment for predicting stock market trends [7, 28, 36]. However, for cryptocurrencies, it tends to have an influence rather than a correlation.

Following these insights, several studies [2, 14, 27] integrated sentiment analysis with LSTM models to predict next-day trading prices. They gathered posts related to cryptocurrencies, assigned each post a sentiment score, and merged these scores with trading data into a singular vector for the LSTM model. Their results revealed a modest enhancement in forecasting outcomes compared to models relying solely on trading data. Furthermore, Huang et al. [14] supported their approach with an autoregressive model that improved accuracy and recall by 18.5%.

However, these studies [2, 8, 14, 20, 27, 30] overlooked the influence of interaction levels on posts’ impact such as likes, comments, shares, and exclusively focused on online communities as fluctuation factors, neglecting other crucial factors [19].

In addition to the findings of Krisoufek [19], the other modalities have proven to be correlated with the stock market, such as search volume [10, 34]

3 APPROACH: COMFORE

To address our first research question (RQ1), this paper proposes a novel strategy for forecasting next-day cryptocurrency volatility by leveraging all indicative factors. To tackle RQ2, we introduce an ensemble learning approach that leverages STL and LSTM architectures, known for their proficiency in learning hidden trends within sequential data, especially for short-term forecasting. As depicted in Figure 1, our approach comprises multiple LSTM weak learners (i.e., forecasters) that collaboratively construct a cryptocurrency volatility prediction model. Each LSTM model \( M \) is trained on a subset sampled from the original dataset and subsequently assigned a weighted score \( w \) based on its performance during the inference phase. Each subset is sampled using the sampling weights \( s \) for the \( N \) samples. Here, a sample denotes the input sequence \( x \) associated with its respective price, where \( k \) is the length of the input layer and \( t \) is the time stamp (i.e., day).

Initially, the sampling weights denoted as \( s \) are set to \( \frac{1}{N} \), indicating that all samples carry equal importance in the learning process. Then, for each LSTM forecaster \( M \), the following steps are performed:

1. Randomly select \( 1 < l \leq L \) features from the total \( L \) features
2. Train the LSTM model \( M \) using the sampled subset. This involves learning the weights \( W_{f.j}, W_{b.j}, W_{c.j}, W_{o.j} \) and biases \( b_{f.j}, b_{i.j}, b_{c.j}, b_{o.j} \) of the LSTM.
3. Calculate the errors \( e \) of the LSTM forecaster \( M \) for each sample \( n \).
4. Compute the total error \( E \) of the \( j \)-th LSTM forecaster, such that:
   \[
   E_j = \sum_{n=1}^{N} s_n^{(j)} e_n^{(j)}
   \]
5. Calculate the weight \( w \) of the \( j \)-th forecaster as follows:
   \[
   w_j = \log \left( 1 - \frac{E_j}{E} \right).
   \]
6. Update the sampling weights \( s_n^{(j)} \) of all samples as follows:
   \[
   s_n^{(j)} = \frac{s_n^{(j)} \exp \left( w_j I \left[ y_n \neq M_j(x_n) \right] \right)}{Z_j},
   \]
   where \( Z_j \) is a normalization factor making \( s_n^{(j)} \) a distribution.

   The normalization factor \( Z_j \) ensures that the updated weights \( s_n^{(j)} \) form a valid probability distribution. The indicator function \( I \left[ y_n \neq M_j(x_n) \right] \) checks if the predicted output of the \( j \)-th model for the \( n \)-th sample does not match the true output for this sample. If true, \( I \) returns 1; otherwise, it returns 0.

   After all LSTM forecasters are trained, we calculate the final ensemble prediction \( \hat{y}_n^{(M)} \) for a given sample \( n \) as follows:
   \[
   \hat{y}_n^{(M)} = \frac{\sum_{j=1}^{J} \hat{y}_n^{(j)} w_j}{\sum_{j=1}^{J} w_j},
   \]
   where \( J \) signifies the total number of LSTM forecasters, \( \hat{y}_n^{(j)} \) represents the prediction made by the \( j \)-th LSTM model \( M_j \) for the \( n \)-th sample, and \( w_j \) denotes the weight of the \( j \)-th LSTM forecaster. This weight reflects the relative contribution of each forecaster’s prediction towards the final ensemble prediction.

   Each LSTM forecaster in our ensemble provides a one-day-ahead forecast of the price of Bitcoin (\( B \)) in US dollars (\( S \)). To do this, the forecaster uses a seven-day window of historical data, denoted as \( x_{t-7} \). Here, \( x_t \) represents a normalized feature vector of length 18 corresponding to day \( t \). This vector combines representations from four distinct modalities: trading data, sentiments, blockchain data, and search volumes. Each of these modalities contributes to the richness and complexity of the feature space and is described in greater detail as follows:

3.1 Input modalities

While there is a multitude of cryptocurrencies available, Bitcoin (\( B \)) remains the most prominent, and data related to it is readily accessible. Therefore, this study centres on Bitcoin as the primary use case. We have gathered a comprehensive dataset spanning from 2012 through the end of 2020, comprising the following modalities:
3.1 Trading data: Trading data encapsulates a sequential collection of attributes that depict the Bitcoin market’s dynamics. Each timestamp $t$ corresponds to a single day, with the opening price recorded at 00:00 and the closing price at 23:59. At each $t$, the following market characteristics are noted:

- **Open Price ($\delta$)**: The price of one Bitcoin (₿) at the beginning of $t$.
- **High Price ($\delta$)**: The peak price of one ₿ reached within the period of $t$.
- **Low Price ($\delta$)**: The lowest price of one ₿ recorded within the period of $t$.
- **Close Price ($\delta$)**: The price of one ₿ at the end of $t$.
- **Bitcoin Volume (₿)**: The total quantity of ₿ traded during the interval of $t$.
- **Dollar Volume ($)**: The total value in dollars of ₿ traded during $t$.
- **Weighted Price ($\delta$)**: The weighted average price of ₿ traded during $t$.
- **Average Transaction Fee ($)**: The average transaction fees charged by the top 20 trading and exchange platforms during $t$.
- **Number of Transactions**: The cumulative number of transactions carried out during $t$.

3.1.2 Public awareness: In today’s world, tweets serve as an instantaneous source of news. Hence, in this study, they are perceived as a crucial indicator of cryptocurrency price trends. Accordingly, we gathered tweets responding to the wildcard queries: “*Bitcoin*” and “*BTC*”, resulting in a corpus of over 120 million tweets. To filter out spam, tweets containing the hashtag “*#Bitcoin*” but not within the main text body were excluded.

The primary goal behind this data collection is to conduct sentiment analysis and incorporate the resulting sentiment scores into our model, thereby encoding the factor of public awareness. To this end, we utilized two widely recognized and publicly available sentiment analysis methodologies: Vader [15] and Deeply Moving [33]. Both techniques yield a sentiment score $\theta_y$ for tweet $y$, ranging from $-1$ (extremely negative) to 1 (extremely positive), while 0 stands for neutral. The dual-method approach was chosen to balance their inherent sensitivity—intensifying the sentiment score when both methods agree and diminishing it when they disagree.

It is important to note that not all cryptocurrency-related tweets have an equal impact on the market. The key determinant here is the level of engagement (i.e., the tweet’s reach and number of interactions). According to recent statistics on Twitter engagement\(^3\), the median number of likes and comments per tweet is zero, implying a substantial portion of tweets receive little to no engagement. Thus, assigning equal sentiment weights to all tweets would not accurately represent the real-world scenario.

To address this, we propose weighting each tweet in a manner that gives higher consideration to those with substantial engagement. Owing to the complexity in determining the relative importance of different engagement factors, we opted for a straightforward approach to compute the weight $\omega_y$ of each tweet $y$ as the harmonic mean of the number of likes, comments, retweets, and quotes. Subsequently, these weights were normalized using min-max normalization. The final sentiment score $\theta_y'$ is then multiplied by $\omega_y$: $\theta_y' = \theta_y \times \omega_y$.

To create a daily sequence of sentiment scores, all weighted sentiments of tweets posted on a given day $t$ are averaged, generating a singular value representing the overall sentiment on Twitter for that day.

3.1.3 Blockchain details: As asserted in the financial study by Kristoufek [19], blockchain metrics significantly influence the price dynamics of cryptocurrencies, notably ₿. As a result, our strategy incorporates several blockchain metrics that have been demonstrated to impact the market in financial research [19]. Relying on various blockchain data sources\(^4\), we include the following indicators for each day $t$:

- **Hash Rate**: An estimation of the computational speed at which the ₿ network is operating.
- **Block Size**: The magnitude of a completed ₿ block.
- **Block Time**: The time consumed to mine and generate a new ₿ block.
- **Network Difficulty**: The complexity associated with mining ₿ blocks across the network.
- **Active Addresses**: The aggregate number of functioning addresses.
- **Mining Profitability**: The projected average profit yielded from mining a single ₿ block.

All the above-listed blockchain attributes follow a sequential pattern with a one-day gap between each recorded data point.

3.1.4 Search volumes. Public curiosity can often be gauged through a variety of behaviours, and in today’s digital age, online searches serve as a significant reflection of societal interest. Thus, we incorporate search volumes into our analysis, specifically focusing on queries involving either the term “Bitcoin” or the acronym “BTC”\(^5\). Given that Google, being utilized by 52% of the global population\(^5\), is one of the most widely used search engines, we have leveraged the Google API to collect search data for the specified terms. This API returns the volume of searches conducted within a particular time span. These search volumes are then aggregated for each day and normalized, providing an indication of public curiosity on a daily basis.

3.2 Fluctuation Analysis

Our primary objective is to augment the decision-making process for cryptocurrency investors. While price prediction is a crucial aspect, it’s insufficient given the varying volatility in the cryptocurrency market. During low market fluctuation, model predictions align closely with the actual value. In contrast, during high volatility, the predicted price can significantly diverge. Existing methodologies do not encapsulate this level of fluctuation. Hence, in addressing RQ3, we aim to analyze these fluctuations using different variations of the model, each trained using a unique combination of modalities and dropout rates.

3.2.1 Input varieties: Given the model described above, and incorporating data from Trading, Twitter Sentiments, Blockchain Details, and Online Search Volumes, we train multiple model variations. Each

---

\(^3\)https://review42.com/resources/google-statistics-and-facts/  
\(^4\)Blockchain.com, YCharts, BitInfoChart, and Nasdaq Data Link  
\(^5\)https://mention.com/en/reports/twitter/engagement/#2
model variation, utilizing different combinations of modalities, generates a slightly distinct price forecast. In terms of architecture, these model variations only differ in the length of their input layer.

- **Trading Data**: The fundamental input for cryptocurrency price forecasting, consisting of eight essential features: ["Open", "High", "Low", "Close", "Volume BTC", "Volume Currency", "Weighted Price", "Average Fees"].
- **Twitter Sentiments**: This variant of our model uses only Twitter sentiments, captured through "Weighted Twitter Sentiments" and "Tweet Volumes".
- **Trading Data and Blockchain Details**: This model variant incorporates trading data combined with blockchain details: ["Hash Rate", "Block Size", "Block Time", "Network Difficulty", "Number of Active Addresses", "Mining Profitability"].
- **Trading Data and Search Volumes**: This model variant merges trading data with ["Online Search Volumes (Google Searches)"] to improve its predictive power.

3.2.2 Model dropouts: To create variability in our model’s predictions and thus estimate an output distribution, we introduce different dropout rates at various layers of the trained model. This method also serves as a regularization technique to counter overfitting. Here are the variants generated with different dropout rates:

- **Models V1, V2, V3**: These models implement dropout rates of 0.1, 0.2, and 0.35 respectively at the last hidden layer of the LSTM.
- **Models V4, V5, V6**: These variants use dropout rates of 0.1, 0.2, and 0.35 respectively at the output of the fully connected layer (fc1).

3.2.3 Predicted Price Distribution: We propose that the distribution of fluctuations can provide valuable insights for decision-makers regarding the most suitable actions to take. Paired with the price forecast, they can offer a measure of the forecast’s uncertainty. To estimate the parameters (i.e., Mean and Variance) of the distribution on the sample of outputs, we utilized Maximum Likelihood Estimation (MLE).

Given the outputs \( O = \{o_1, o_2, \ldots, o_10\} \) from the ten model variants, the aim is to estimate the parameter set \( \hat{\theta} = \{\mu, \sigma^2\} \) that maximizes the likelihood function \( L(\theta; O) \). This can be represented as:

\[
L(\theta; O) = \prod_{i=1}^{10} N(o_i; \theta)
\]

Then, the set of parameters that maximizes this likelihood function is given by: \( \hat{\theta} = \arg \max L(\theta; O) \).

Here, \( N(o_i; \theta) \) denotes the normal distribution of the output \( o_i \) with the parameter set \( \theta \), and \( \hat{\theta} \) represents the estimated parameters that maximize the likelihood function.

4 EXPERIMENTS

In this section, we assess CoMForE by comparing its performance to several other baseline approaches. To ensure reproducibility and facilitate further exploration, we have made our implementation available on GitHub.\(^6\)

4.1 Experimental Setup

Across all experiments, we maintain the same setup and parameters. These are as follows:

- Loss function during training: Mean Squared Error (MSE)
- Optimizer: *Adam*, Initial learning rate: 0.0003 and Number of epochs: 200
- The training process takes place on a GPU server with the following specifications: an AMD Ryzen Threadripper 1950X 16-Core Processor, 128GiB of system memory, and an NVIDIA GV100 GPU.
- All experiments employ the Mean Absolute Error (MAE) as the validation loss, calculated every ten epochs.

4.2 Datasets

This study focuses on Bitcoin (₿), the dominant cryptocurrency with readily accessible data. We gathered Bitcoin data from 2016 to 2020, yielding 1825 data points (days). The data was split into 70% for training, 15% for validation, and 15% for testing, as shown in Figure 2. Next, we proceed to a detailed exploration of each data modality.

\[\text{Figure 2: Data split for training, validation, and testing (70% training, 15% validation, 15% testing)}\]

4.2.1 Trading Data: Acquiring comprehensive and accurate historical trading data for cryptocurrencies can be a challenging task due to the presence of missing values and occasional skipped days in many data sources. To address this, we have amalgamated data from several sources, each supplementing the gaps in the other. We also account for the potential variances in values across exchanges, given the lack of a standardized pricing protocol for cryptocurrencies. The sources for our data are as follows:

- **Kaggle**: This open-source dataset comprises 1-minute interval data of $\hat{p}$ prices, collected from January 1, 2012, to March 1, 2021.
- **Binance**: As a major cryptocurrency exchange, Binance publicly provides its data. In the absence of an API, data was manually extracted from the website.
- **Coinbase**: Unlike Binance, Coinbase offers APIs to gather trading data. As a major cryptocurrency exchange, its data is widely used, including by organizations like Google.

4.2.2 Public Awareness: For sentiment analysis of the collected tweets, a distinct subdiscipline, this paper employs two renowned,
pre-trained sentiment analysis tools: Vader[15] and Deeply Moving[33].

Vader [15], a rule-based model tailored for social media sentiment analysis, has exhibited superior results in recent evaluations compared to various rule-based or machine learning approaches. It is expected to generalize better across different contexts. The model’s output helps classify a tweet into one of nine sentiment classes, ranging from extremely negative to extremely positive.

Deeply Moving [33] is another tool that examines text as a holistic entity, maintaining word correlations. This method is primarily designed for movie reviews sentiment analysis, meaning that the text is formally represented, which is different from the structure used in Twitter’s tweets, such as emojis, hashtags, and abbreviations. Despite the difficulty of fine-tuning this model on tweet sentiments due to the scarcity of labelled data, we aim to utilize both models to harness their unique strengths - Vader’s explicit training on tweets and Deeply Moving’s ability to comprehend sentiments of an entire sentence.

Both sentiment analysis models (Vader, Deeply Moving) are applied to all previously gathered tweets. The final sentiment score is then derived as the average of the scores from both models.

4.2.3 Blockchain data: As highlighted in Section 3.1.3, the blockchain data has been sourced from various providers.

4.2.4 Search volumes: As previously stated in Section 3.1.4, we employed Google’s search engine statistics via its API in this study. Figure 3 represents the search volumes associated with the query "Bitcoin" from 2013 to 2021.

4.3 Baselines

To evaluate CoMForE, we selected six notable and modern techniques, each with a unique architecture or different data types, to forecast Bitcoin prices. These include:

- ARIMA19 [4]: Utilizes the ARIMA model, focusing exclusively on Bitcoin trading data to forecast the next day’s prices.
- BNN17 [16]: Employs Bayesian Neural Networks (BNN) to predict future Bitcoin prices, harnessing trading data.
- LSTM20A [35]: Adopts a multivariate LSTM model tailored for cryptocurrency price predictions.
- LSTM20B [26]: Uses an LSTM architecture, grounded on trading data, to predict price trajectories.
- GRU20 [11]: Employs a Gated Recurrent Unit (GRU) infused with recurrent dropout to enhance Bitcoin price prediction accuracy.
- AdaBoost21 [8]: Leverages the traditional AdaBoost algorithm, incorporating multiple decision tree weak learners for its predictions. Its straightforward design achieves noteworthy results, even surpassing more sophisticated methods.

Since the source codes of these approaches are not publically available, we re-implemented them from scratch and reproduce their results. To ensure a fair comparison, we evaluated all the baseline methods, including CoMForE, on the exact same dataset within the same time period.

4.4 Results and Discussion

4.4.1 Modality Analysis. To address RQ1, we conducted an evaluation using both CoMForE and an LSTM-based architecture to investigate the effectiveness of various combinations of data modalities. Both models underwent training for 200 epochs and adhered to the experimental setup outlined earlier. Note that the LSTM model shares the same architecture as a single weak learner in CoMForE, but trained with the full dataset. The results, displayed in Table 1, present the performance of both models using (a) only trading data, (b) only sentiments, and (c) a combination of trading data and sentiments. Remarkably, the results from the sentiment data alone (b) underscore the correlation between social media sentiment and cryptocurrency prices. The results derived from the combination of trading data and Twitter sentiments (c) affirm that this combination yields the most accurate results. This suggests that social media sentiments can effectively complement trading data in capturing cryptocurrency price fluctuations, thereby enhancing the model’s forecasting capability. Specifically, adding sentiment analysis to both models can improve the results by up to 15% in comparison to using only trading data.

Additionally, we performed further experiments applying LSTM and CoMForE on trading data, supplemented with various other modalities; (d) the hash rate, (e) search volume, and (f) blockchain data encompassing all six modalities. The evaluation results of both models are presented in Table 2. While the error for (d) is marginally higher for CoMForE, which is possibly due to the boosting process, the lower error observed for LSTM when incorporating the hash rate suggests it as a beneficial factor for a more precise price forecasting. The findings from (e) suggest that search volumes did not notably influence the performance of either LSTM or CoMForE when combined with trading data. However, it can be speculated that search volumes could exhibit enhanced performance when integrated with other modalities, such as social media sentiments, as both represent a measure of “public awareness”.

The outcomes presented in column (f) for both models substantiate that incorporating blockchain data and trading data significantly enhances predictive accuracy compared to using the hash rate alone. However, the testing errors encountered with the CoMForE are notably larger than those found with the LSTM model. As demonstrated in Figure 4, when blockchain data was utilized as the sole input for both models, CoMForE didn’t precisely forecast the price, though it could track the price trend. These results highlight a noticeable correlation between blockchain data and cryptocurrency prices, which can assist in the prediction of cryptocurrency prices.
To the best of our knowledge, this is the first study to integrate blockchain data into an analytical framework for predicting cryptocurrency prices. For further insights into the interplay between different modality combinations, Figure 5 showcases the qualitative results of CoMForE and LSTM over a randomly sampled 100-day testing interval across all combinations.

![Figure 5: CoMForE and LSTM outcomes across various modality combinations over 100-day interval.](image)

In our comprehensive evaluation, Table 3 presents the cumulative impact of combining all four modalities: trading data, social media sentiments, search volumes, and blockchain data. As clearly illustrated, this combination significantly enhances the results, indicating their complementary nature. Specifically, we observed an improvement of $75.312 in MAE, which corresponds to a 36.32% enhancement in forecasting accuracy. To visually corroborate this, Figure 6 offers a qualitative comparison over a span of 100 days, where the forecasted trajectories from both models closely align with the actual price trends.

![Figure 6: Qualitative comparison of LSTM and CoMForE predictions over 100 days.](image)

In response to RQ2, Tables 1, 2, and 3 reveal the anticipated outcome, where CoMForE shows substantial superiority over the LSTM in all scenarios. In cases (a) and (c), the LSTM’s overall validation error appears lower, which could be due to the inconsistent trends.
Zeyd Boukhez, Azeddine Bouabdellah, Cong Yang, & Jan Jürjens

Figure 6: LSTM and CoMForE forecasting results with all modalities over 100-day interval.

and volatilities in price over time. Since LSTM typically produces smoother forecasts than the weak learners of CoMForE, it yields a lower error during periods of high volatility. As depicted in Figure 2, periods from 2012 to 2017 and the end of 2018 to 2019 exhibited minor fluctuations compared to other timeframes.

Figure 7 presents the qualitative outcomes of the CoMForE and LSTM models over a randomly selected 100-day testing interval for combinations (a), (b), and (c). Similarly, Figure 8 illustrates the results for combinations (d), (e), and (f). In both figures, the blue curve represents the actual price. Meanwhile, the red and green curves depict the price forecasts generated by the LSTM and CoMForE models, respectively. Notably, CoMForE consistently outperforms the single learner, LSTM, especially for the combinations (a), (b), and (c). Although the forecasting error in (b) is notably higher than when only using trading data, it is significant to acknowledge that CoMForE, despite being fed with a solitary input per day (i.e., Twitter sentiments), managed to detect the relationship between prices and sentiments, subsequently producing a meaningful price prediction. As illustrated in Figure 7b, CoMForE successfully tracks the price trend using sentiment data alone. In contrast, the LSTM model has difficulty learning patterns from the input data and tends to produce almost constant forecasts over time.

To gain a deeper understanding of the model’s performance, we computed the MAE distribution for the entire testing set as illustrated in Figure 9. As can be seen, the forecasted price deviates by $\pm 500$ from the actual price of 1 Bitcoin approximately 68% of the time, yielding to an average error of 3.02%. This fluctuation is particularly significant given that the average price within the testing set is $16553.59. This deviation emphasizes the commendable precision of the model’s forecasting relative to the actual price.

4.4.2 Comparison against baseline approaches. Additionally, we evaluate the performance of CoMForE, trained with all modalities, in comparison to baseline methodologies. The comparison is conducted over the same dataset, within the period from January 1, 2020, to July 1, 2020. It is noteworthy that all models in this comparison are forecasting Bitcoin prices for the subsequent 24 hours. As delineated by the results in Table 4, it is evident that CoMForE outperforms the other models, which capitalizes on multiple data modalities and on the power of ensemble learning, enhancing the understanding of the underlying trends and thereby improving the accuracy of its predictions.

4.4.3 Fluctuation Analysis. In response to RQ3, we propose to equip the investor with not only the price forecast but also the range of potential price fluctuations. Given the complexity of quantifying this, we present in Figure 10 a set of outcomes obtained using the approach described in Section 3.2 over four random days. As illustrated, the certainty of price forecast can vary day-to-day. On days marked by significant volatility, the price distribution appears flat and exhibits a high standard deviation, indicating a high likelihood of deviation from the model’s singular prediction. Conversely, on days when the model predicts with a high degree of certainty, the price distribution tends to peak sharply and display a lower standard deviation. These prediction distributions serve as a valuable resource for investors as they offer clear insights into the model’s confidence in its predictions for a specific day.

4.4.4 Long-term forecasting. In light of the promising obtained results, we further explored the limitations and performance of
Table 4: Evaluation Comparison Between the results of CoMForE and baseline approaches on test data.

|                | CoMForE | BNN17 | LSTM20A | LSTM20B | AdaBoost21 | GRU21 | ARIMA19 | GRU20 |
|----------------|---------|-------|---------|---------|------------|-------|---------|-------|
| RMSE ($)       | 158.929 | 221.642 | 166.875 | 202.787 | 197.544 | 208.006 | 584.384 | 185.154 |
| MAE ($)        | 132.027 | 184.124 | 148.628 | 177.02  | 156.346  | 160.44  | 542.73  | 157.632 |

Figure 8: Visual representation of the forecasting performance of LSTM and CoMForE over a 100-day interval, tested using trading data combined with (d) the hash rate, (e) search volume, and (f) blockchain data.

Figure 10: Predicted price distribution of four randomly selected days. Mean denotes the predicted price.

ComForE and multimodality combination in generating long-term predictions. Therefore, we continuously predicted the price by appending the last prediction to the input of the prediction of the subsequent day. This process is reiterated for up to the 30th day. It is crucial to note that only historical trading data is considered since other data modalities cannot be predicted. This assumption implies a likely performance dip because the model is model operating with less data for each successive day’s prediction. Figure 11 showcases the results of this experiment, with the red curve representing the mean MAE of each time-window prediction. The boundaries of the blue area, meanwhile, indicate the highest and lowest MAE. As anticipated, the predictive accuracy is at its peak on the first day and progressively diminishes for longer-term forecasts.

Figure 11: Bitcoin price prediction absolute testing errors over different time windows.
5 CONCLUSION

In this paper, we explored the impact of various independent factors on cryptocurrency volatility through a comprehensive multimodal forecasting approach, aiming to provide investors with accurate short-term price forecasts based on correlated sources (RQ1). The proposed approach harnesses the power of ensemble learning, combining multiple weak LSTM learners (RQ2). Understanding that investors need a more comprehensive view of the market, we introduced the concept of fluctuation distributions, which offer a broader perspective on the market and provide insights into the reliability of the price forecast (RQ3).

For future work, we plan to expand the application of this approach to other cryptocurrencies, incorporating the Bitcoin price as a vital variable. We also propose to generate and train the model on synthetic trading data characterized by extremely high volatility. This approach stems from our hypothesis that preparing the model for unprecedented market volatility will enhance its adaptability to future surprises inherent in the cryptocurrency markets.

REFERENCES

[1] David Aboody, Omri Even-Tov, Reuven Levyhav, and Brett Trueman. 2018. Overnight returns and firm-specific investor sentiment. Journal of Financial and Quantitative Analysis 53, 2 (2018), 485–505.
[2] Jethin Abraham, Daniel Higdon, John Nelson, and Juan Ibarra. 2018. Cryptocurrency price prediction using tweet volumes and sentiment analysis. SMU Data Science Review 1, 3 (2018), 1.
[3] M. Eren Akbiyik, Mert Erkul, Killian Kampf, Kaiva Vasiliauskaite, and Nino Antulov-Fantulin. 2023. Ask ‘who’, not ‘what’. Bitcoin volatility forecasting with twitter data. In Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. 688–696.
[4] Saad Ali Almahari. 2019. Using machine learning ARIMA to predict the price of cryptocurrencies. The ISC International Journal of Information Security 11, 3 (2019), 139–144.
[5] Rayner Alfred, Joe Henry Obit, Mohd Hanafi Ahmad Hijazi, Ag Asri Ag Ibrahim, et al. 2015. A performance comparison of statistical and machine learning techniques in learning time series data. Advanced Science Letters 21, 10 (2015), 3037–3041.
[6] Reem K Alkhodhairy, Shahad R Aljalhami, Norah K Rasayni, Jawharah F Alshoalabil, Amal A Al-Shargabi, and Abdulsalat Abulaboud. 2021. Bitcoin Candlestick Prediction with Deep Neural Networks Based on Real Time Data. CMC- COMPUTERS MATERIALS & CONTINUUM 68, 3 (2021), 3215–3233.
[7] Johan Bollen, Huana Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. Journal of computational science 2, 1 (2011), 1–8.
[8] Julien Chevallier, Dominique Guégan, and Stéphane Goutte. 2021. Is It Possible to Forecast the Price of Bitcoin? Forecasting 3, 2 (2021), 377–420.
[9] David LEE Kuo Chuen, Li Guo, and Yu Wang. 2017. Cryptocurrency: A new investment opportunity? The Journal of Alternative Investments 20, 3 (2017), 16–40.
[10] Zhi Du, Joseph Engberg, and Pengjie Gao. 2011. In search of attention. The journal of finance 66, 3 (2011), 1461–1499.
[11] Aniruddha Dutta, Sakti Kumar, and Mehezi Basu. 2020. A gated recurrent unit approach to bitcoin price prediction. Journal of Risk and Financial Management 13, 2 (2020), 23.
[12] James V Hansen, James B McDonald, and Ray D Nelson. 1999. Time series prediction with Genetic-Algorithm designed neural networks: An empirical comparison with modern statistical models. Computational Intelligence 15, 3 (1999), 171–184.
[13] Mohammad Asif Husain, Rezaul Karim, Ruppa Thulasiram, Neil DB Bruce, and Yang Wang. 2018. Hybrid deep learning model for stock price prediction. In 2018 ieee symposium series on computational intelligence (ssci). IEEE, 1837–1844.
[14] Xin Huang, Wenbin Zhang, Yiyi Huang, Xunjiao Tang, Mingli Zhang, Jayachan-
der Suryibyla, Vasileios Iosifidis, Zhen Liu, and Ji Zhang. 2021. LSTM Based Sensitivity Analysis for Cryptocurrency Prediction. arXiv preprint arXiv:2105.14804 (2021).
[15] Marton Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the International AAAI Conference on Web and Social Media. Vol. 8.
[16] Huius Jang and Jaewook Lee. 2017. An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. IEEE Access 6 (2017), 5427–5437.

[17] Kaustubh Khare, Omkar Darekar, Prafulla Gupta, and VZ Attar. 2017. Short term stock price prediction using deep learning. In 2017 2nd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT). IEEE, 482–486.
[18] Young Bin Kim, Jun Gi Kim, Wook Kim, Jae Ho Im, Tae Hyeong Kim, Shin Jin Kang, and Chang Hun Kim. 2016. Predicting fluctuations in cryptocurrency transactions based on user comments and replies. PloS one 11, 8 (2016), e0161197.
[19] Ladaslav Kristoufek. 2015. What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. PloS one 10, 4 (2015), e0123923.
[20] Deepak Kumar and SK Rath. 2020. Predicting the trends of price for etherum using deep learning techniques. In Artificial Intelligence and Evolutionary Compu-
tations in Engineering Systems. Springer, 103–114.
[21] Connor Lamon, Eric Nielsen, and Eric Redondo. 2017. Cryptocurrency price prediction using news and social media sentiment. SMU Data Sci. Rev 1, 3 (2017), 1–22.
[22] Chang Li, Dongjin Song, and Dacheng Tao. 2019. Multi-task recurrent neural networks and higher-order Markov random fields for stock price movement prediction. Multi-task RNN and higher-order MRFs for stock price classification. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. 1141–1151.
[23] Ioannis E Livieris, Emmanuel Pintelas, Stavros Stavroyiannakis, and Panagiotis Pintelas. 2020. Ensemble deep learning models for forecasting cryptocurrency time-series. Algorithms 13, 5 (2020), 121.
[24] Isaac Madan, Shaurya Saluja, and Anjia Zhao. 2015. Automated bitcoin trading via machine learning algorithms. URL: http://cs229. stanford. edu/proj2014/Isaac%20Madan 20 (2015).
[25] Sean McNally, Jason Roche, and Simon Caton. 2018. Predicting the price of bitcoin using machine learning. In 2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP). IEEE, 339–343.
[26] Mohammed Moadassir, Shada Bembaiva, Devrim Unal, and Mohammad Ham-
douche. 2020. Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. Neuro Computing and Applications (2020), 1–15.
[27] Sergey Nasekin and Cathy Yi-Hsuan Chen. 2019. Deep learning-based cryptocurrency sentiment construction. Available at SSRN 3107874 (2019).
[28] Thien Hui Nguyen and Kiyoaki Shirai. 2015. Topic modeling based sentiment analysis on social media for stock market prediction. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 1354–1364.
[29] Mahla Nikou, Gholamreza Mansourfar, and Jamshid Bagherzadeh. 2019. Stock price prediction using deep learning algorithm and its comparison with machine learning algorithms. Intelligent Systems in Accounting, Finance and Management 26, 4 (2019), 164–174.
[30] Emmanuel Pintelas, Ioannis E Livieris, Stavros Stavroyiannakis, Theodore Kot-
silieris, and Panagiotis Pintelas. 2020. Investigating the problem of cryptocurrency price prediction: a deep learning approach. In IFIP International Conference on Artificial Intelligence Applications and Innovations. Springer, 99–110.
[31] Christophe Schineckus. 2020. The good, the bad and the ugly: An overview of the sustainability of blockchain technology. Energy Research & Social Science 69 (2020), 101614.
[32] Sreelekshmy Selvin, R Vinayakumar, EA Gopalakrishnan, Vijay Krishna Menon, and KP Soman. 2017. Stock price prediction using LSTM, RNN and CNN-sliding window model. In 2017 international conference on advances in computing, communications and informatics (icacci). IEEE, 1643–1647.
[33] Reem R Alkhodhair, Mohammad Asif Husain, Rezaul Karim, Ruppa Thulasiram, Neil DB Bruce, and Yang Wang. 2018. Hybrid deep learning model for stock price prediction. In 2018 ieee symposium series on computational intelligence (ssci). IEEE, 1837–1844.
[34] Xun Huang, Wenbin Zhang, Yiyi Huang, Xunjiao Tang, Mingli Zhang, Jayachander Suryibyla, Vasileios Iosifidis, Zhen Liu, and Ji Zhang. 2021. LSTM Based Sensitivity Analysis for Cryptocurrency Prediction. arXiv preprint arXiv:2105.14804 (2021).
[35] Zeyd Boukhers, Azeddine Bouabdellah, Cong Yang, & Jan Jürjens Zeyd Boukhers, Azeddine Bouabdellah, Cong Yang, & Jan Jürjens Zeyd Boukhers, Azeddine Bouabdellah, Cong Yang, & Jan Jürjens