Computer Science
&
Data Science

Capstone Report - Fall 2024

Predicting Bitcoin Price Movements Using Sentiment Analysis: A Data-Driven Approach Integrating Social Media Indicators and Financial Metrics

Alibi Nauanov,
Douaa Zouhir,
Jalal Haider

supervised by
Xianbin Gu

Preface

The ever-evolving landscape of financial markets, particularly the volatile cryptocurrency sector, has spurred our deep interest in leveraging advanced data analysis techniques to predict market trends. With a background in computer science and a passion for financial technologies, we embarked on this capstone project to explore the intricate relationship between public sentiment and Bitcoin price movements. Inspired by the significant impact that social media sentiment can have on cryptocurrency markets, this project aims to develop a predictive model that integrates sentiment analysis with financial indicators. The target audience for this report includes academic researchers, financial analysts, and cryptocurrency enthusiasts who seek innovative methods to enhance market prediction accuracy. This project is crucial as it provides insights into how sentiment-driven data can be harnessed to navigate the unpredictable nature of cryptocurrency investments.

Acknowledgements

We would like to extend our deepest gratitude to our capstone supervisor, Xianbin Gu, for his exceptional guidance and unwavering support throughout this project. Professor Gu's expertise in machine learning models and his insightful advice on the research methodology were instrumental in shaping the direction and success of this study. His patience and willingness to provide detailed feedback significantly enhanced the quality of our work and deepened my understanding of complex analytical techniques.

Abstract

Bitcoin's price volatility presents a major challenge for investors and analysts aiming to predict market trends. The unpredictable nature of cryptocurrency markets, driven by both economic factors and public sentiment, makes forecasting particularly complex. To address this, we developed an ensemble-based machine learning model that integrates sentiment analysis from tools like VADER, TextBlob, and BERT with financial indicators such as price fluctuations and trading volume. By combining the strengths of multiple sentiment analysis models, our approach enhances predictive accuracy and reliability. Evaluation results indicate that the ensemble model significantly outperforms individual models, offering improved accuracy in forecasting Bitcoin price movements. This study demonstrates the value of integrating sentiment analysis with financial metrics for more dependable cryptocurrency market predictions.

Keywords

Bitcoin, Sentiment Analysis, Machine Learning, Cryptocurrency Prediction, VADER, BERT, TextBlob, Random Forest

Contents

1 Introduction

The cryptocurrency market, with Bitcoin at its forefront, is renowned for its extreme volatility and susceptibility to a myriad of influencing factors, including public sentiment. Traditional financial models often fall short in predicting Bitcoin price movements due to the market's inherent unpredictability and the significant impact of social media sentiment. This capstone project aims to bridge this gap by developing a predictive model that integrates sentiment analysis from platforms like Twitter with key financial metrics such as price fluctuation, trading volume, and market dominance. By analyzing historical sentiment data alongside market trends, the project seeks to create an ensemble-based machine learning model capable of forecasting Bitcoin price movements with enhanced accuracy. The achievement of this objective not only contributes to academic research but also offers practical tools for investors and traders navigating the volatile cryptocurrency landscape.

2 Related Work

The prediction of Bitcoin prices, given the market's volatility and sensitivity to public perception, has attracted significant research attention. Existing literature on Bitcoin price forecasting can be broadly divided into three complementary areas: sentiment analysis approaches, ensemble modeling techniques, and feature extraction methods. While these studies have individually demonstrated improvements in predictive accuracy, none have comprehensively integrated sentiment analysis, financial indicators, and ensemble methods into a unified framework that captures the multifaceted nature of Bitcoin price dynamics. Our work positions itself at this intersection, offering a novel synthesis that surpasses the limitations of prior research.

2.1 Sentiment Analysis for Bitcoin Price Prediction

Early studies established that textual data from social media and online forums can be a powerful proxy for investor sentiment. Sattarov et al. [1] applied the VADER sentiment tool to Bitcoin-related news headlines and integrated sentiment scores with asset-specific variables (e.g., price, volume, and market capitalization). Their use of an SVM model outperformed Linear Discriminant Analysis, achieving a 58.5% accuracy rate in predicting price movements. This highlighted the importance of coupling sentiment with traditional financial metrics. Similarly, Critien et al. [2] leveraged Twitter sentiment through VADER and employed neural networks (LSTM, CNN, BiL-

STM) to predict both price direction and magnitude, achieving a notable 77.2% accuracy using a voting classifier.

Other researchers have deepened this exploration. Georgoula et al. [3] demonstrated that positive Twitter sentiment correlates with rising Bitcoin prices, while macroeconomic indicators such as the USD/EUR exchange rate exert a moderating effect. Subsequent works by Gurrib and Kamalov [4] and Passalis et al. [5] confirmed that combining sentiment data with economic and financial information enhances forecast accuracy. Pant et al. [6] and Serafini et al. [7] advanced the field further by utilizing recurrent models (LSTM, ARIMAX) and showing that Twitter sentiment can be a robust predictor of short-term price fluctuations. These studies underscore that while sentiment alone is informative, its full predictive potential emerges when paired with complementary data streams.

Despite these advances, most sentiment-focused studies have centered on a limited set of models or relied heavily on one type of sentiment extraction tool. Furthermore, incorporating multiple sentiment analysis techniques—such as VADER, TextBlob, and BERT—into a single predictive pipeline remains underexplored. Our approach addresses this gap by unifying multiple sentiment extraction methods, thus harnessing their varied strengths and mitigating individual weaknesses.

2.2 Ensemble Methods

Given the complexity and non-stationarity of cryptocurrency markets, relying on a single predictive model often leads to suboptimal results. Ensemble methods have shown promise in this regard. Bâra and Oprea [8] demonstrated that decision tree-based ensembles, like Random Forest (RF), XGBoost, and LightGBM, are effective at handling non-linearities and improving forecast accuracy. Rather [9] and Sin and Wang [10] introduced neural network-based ensembles, including DNN-SVR-Decision Tree combinations and GASEN (a Genetic Algorithm-based Selective Neural Network Ensemble), which carefully select diverse learners to enhance predictive robustness.

Stacking ensembles have further refined these ideas. Ye et al. [11] and Shin et al. [12] illustrated how combining LSTM and GRU models can capture both short- and long-term dependencies in Bitcoin price trends. Livieris et al. [13] integrated LSTM, Bi-LSTM, and CNNs into a deep learning ensemble to reduce variance and bias, while Silva et al. [14] leveraged Variational Mode Decomposition (VMD) and a meta-learner to improve multi-step ahead forecasts.

These ensemble-based strategies validate the idea that integrating multiple models outperforms individual learners. However, prior work has not fully leveraged the synergy between ensemble methods and sentiment-based features. Our work bridges this gap by applying Random Forest ensembles not only to traditional financial indicators and a single sentiment score but to a diverse set of sentiment analysis outputs, thereby capturing a richer representation of market mood and behavioral cues.

2.3 Feature Extraction

Beyond model selection, the quality of input features profoundly influences predictive performance. Feature extraction strategies aim to isolate the most relevant factors affecting Bitcoin prices. Htun et al. [15] and Singh et al. [16] examined various selection and transformation techniques, noting that Random Forest, PCA, and Autoencoders help refine input datasets. Deep learning-based methods, such as CNNs [17] and hybrid models combining dandelion optimization with 3D-CNN-GRU [18], have successfully extracted intricate patterns from high-dimensional time-series data.

Wang et al. [19] and Chen [20] emphasized the importance of incorporating lagged explanatory variables and multi-scale non-linear feature extraction for reducing uncertainty. Gupta and Nalavade [21] extended this line of inquiry by using metaheuristics to integrate multiple feature extraction methods, ultimately boosting predictive accuracy. While feature extraction is well studied, the literature primarily focuses on either financial metrics or a single type of sentiment measure. Our approach integrates sentiment data from multiple models (VADER, TextBlob, BERT) with key financial features, ensuring that all relevant predictors—textual and numeric—are meaningfully represented.

2.4 Positioning Our Work

Collectively, the existing literature demonstrates that sentiment analysis, ensemble modeling, and feature extraction techniques can each enhance Bitcoin price predictions. Yet, these advances have largely occurred in isolation. Past studies have focused on either using a single sentiment model or improving a particular predictive modeling technique or feature extraction method.

Our work distinguishes itself by holistically integrating these three strands of research. We unify multiple sentiment analysis tools to capture varying dimensions of public mood, apply an ensemble approach (Random Forest) to combine these sentiment scores with rich financial indicators, and adopt advanced feature extraction methodologies to refine the input space. By combining the most promising elements of prior studies—robust sentiment extraction, ensemble modeling, and intelligent feature selection—we aim to achieve more accurate and stable Bitcoin price predictions than any individual technique could provide.

In summary, while prior work has established the importance of sentiment data, validated the effectiveness of ensemble methods, and proven the value of sophisticated feature extraction, our research uniquely synthesizes these components into an integrated predictive framework. This comprehensive approach addresses existing gaps and sets the stage for more reliable Bitcoin price forecasting models.

3 Solution

3.1 Methodology

Our methodology was divided into four main areas, data collection from the aforementioned sources, applying and running the tweets on the the three Sentiment Analysis Models, namely VADER, BERT and TEXTBLOB. Next, we included the combination and the merging of these scores and the prices. Lastly, we incorporated the Model Training by running the merged dataset into the Long Short Term Memory Model (LSTM) for each dataset of the prices and their subsequent models scores. For the three models, we also created an ensemble score which combined the scores of all the three models. In total we generated four different models and visualised the predicted prices against the actual price for a more enhanced understanding of the most useful model.

3.2 Data Collection

For the Bitcoin price sentiment analysis project, the data collection process involved gathering a substantial dataset of tweets to analyze public sentiment. Approximately 285,000 tweets and their corresponding bitcoin price data were sourced from two primary platforms: Hugging Face and Kaggle. The datasets from these sources were merged to create a unified corpus for analysis. This integration resulted in a comprehensive time-series dataset spanning from January 31, 2023, to June 6, 2023. Each tweet was timestamped, enabling chronological organization and alignment with corresponding Bitcoin price data, including values such as the open, close, high, and low prices of Bitcoin during this period. The collected data provides a robust foundation for exploring the relationship between public sentiment and Bitcoin price movements over the specified timeframe.

3.3 Sentiment Analysis Scoring

The analysis incorporated three sentiment models: TextBlob, VADER, and BERT, each offering unique strengths in sentiment classification and interpretation.

TextBlob, a simple yet powerful natural language processing library, was utilized to provide initial sentiment insights. It calculates two key metrics: polarity and subjectivity. Polarity scores range from -1 (negative) to +1 (positive), indicating the sentiment's positivity or negativity. Subjectivity scores range from 0 (objective) to 1 (subjective), reflecting the degree of opinion or factual content in the text. TextBlob's straightforward implementation and efficient performance make it a popular choice for sentiment analysis tasks. Its lexicon-based approach, although not as context-aware as deep learning models, serves as a reliable baseline for sentiment analysis, especially for datasets with diverse linguistic expressions.

VADER, or Valence Aware Dictionary and sEntiment Reasoner, was employed for its specialization in sentiment analysis of short, informal texts such as tweets. Unlike traditional sentiment models, VADER uses a lexicon-based approach to identify the sentiment intensity of words and phrases, making it highly effective for analyzing slang, abbreviations, and emoticons commonly found in social media posts. By assigning weighted sentiment scores to each word and aggregating these scores, VADER delivered an effective sentiment intensity for the tweets, enriching the overall scope of the analysis.

The BERT (Bidirectional Encoder Representations from Transformers) model added a layer of sophistication to the sentiment analysis process. Fine-tuned using a cased tokenizer, BERT leveraged its deep learning architecture to understand contextual relationships within the text, going beyond surface-level sentiment indicators. This model employed pre-trained embeddings and classifier layers to categorize tweets as positive, neutral, or negative, capturing nuanced sentiment variations often overlooked by simpler models. By integrating the outputs of TextBlob, VADER, and BERT, a comprehensive sentiment score was derived for each tweet, reflecting its potential influence on market sentiment and laying a strong foundation for predictive analysis.

3.4 Feature Engineering

To enhance the predictive accuracy of the Bitcoin price movement model, a meticulous and comprehensive feature engineering process was implemented. This stage was crucial in preparing the data to reflect meaningful relationships between social sentiment and market dynamics. It involved integrating the datasets comprising sentiment scores derived from tweets and Bitcoin price data, ensuring synchronization based on the corresponding dates. This alignment provided a structured dataset capable of capturing the nuanced interplay between public sentiment trends and Bitcoin's market behavior, a critical factor in forecasting volatile price movements.

The process as previously mentioned, began with the extraction of sentiment scores from approximately 285,000 tweets collected during the study's timeframe, spanning January 31, 2023, to June 6, 2023. These tweets were analyzed using advanced sentiment analysis tools, including VADER, TextBlob, and BERT, to compute key sentiment metrics such as polarity, which measures positive or negative sentiment, and subjectivity, which gauges the extent of personal opinions expressed in the text. The sentiment outputs were categorized into distinct classes: positive, neutral, negative and compound for the VADER and BERT models, whilst for TextBlob, the classes were distinguished by their polarity and subjectivity scores. These scores served as a primary feature set representing the broader sentiment trends surrounding Bitcoin on social media platforms.

Concurrently, detailed financial data was gathered, encompassing Bitcoin's daily open prices, high prices, low prices, closing prices, and trading volumes. These metrics provided critical insights into market conditions on a day-to-day basis. To create a unified dataset, the financial metrics (high and low prices) and sentiment scores (compound scores from VADER and BERT and polarity/subjectivity scores for TextBlob) were merged accordingly based on their corresponding dates. This alignment ensured that for each day in the dataset, there was a comprehensive record combining sentiment data with financial performance metrics.

After merging, the combined dataset underwent a thorough preprocessing phase. Numerical values, such as sentiment scores and price metrics, were normalized to bring all input features onto consistent scales. This step reduced the risk of dominance by variables with larger magnitudes and improved the stability and speed of the learning process during model training. Moreover, missing values, if any, were carefully handled to avoid introducing noise or bias into the model.

In addition to normalization, temporal sequencing was applied to the dataset. The input features were formatted into sequences of ten consecutive days, capturing temporal dependencies across both sentiment and financial data. These sequences served as input for the Long Short-Term Memory (LSTM) network, which is designed to model time-series data effectively. The use of temporal sequences enabled the model to learn patterns over time, identifying how changes in sentiment and market trends influenced subsequent Bitcoin price movements.

The resulting feature set provided a holistic view of Bitcoin's market behavior, integrating sentiment-driven factors such as public opinion and sentiment polarity with data-driven factors like trading activity and price fluctuations. By capturing both types of information, the dataset positioned the prediction model to analyze complex temporal patterns, ensuring robust performance in a highly volatile and dynamic market environment. This feature engineering process laid a strong foundation for developing a predictive model capable of delivering actionable insights into Bitcoin price movements.

3.5 Model Training

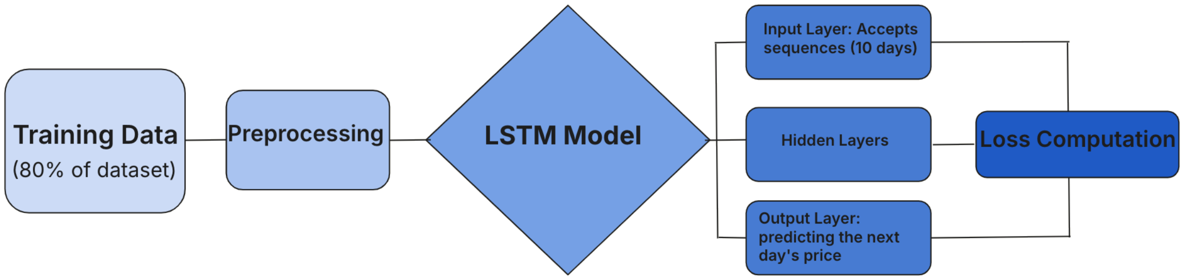

Figure 1: training process

The training process for the prediction model utilized 80% of the dataset, ensuring a balanced representation of various market conditions and sentiment patterns. This approach enabled the model to generalize effectively across different scenarios, capturing the nuanced interactions between market trends and public sentiment. The pipeline began with preprocessing the training data, where input sequences were carefully constructed to include ten days' worth of data. These sequences combined sentiment scores derived from the aforementioned advanced natural language processing techniques, such as VADER, BERT, and TextBlob, with financial metrics such as the open and close prices of the Bitcoin. The inputs were normalized to ensure consistent scales, reducing the risk of bias caused by disproportionate numerical values and enhancing the convergence rate during training.

The formatted data was fed into a Long Short-Term Memory (LSTM) network architecture, specifically designed to handle time-series prediction tasks due to its ability to retain long-term dependencies in sequential data. The architecture consisted of an input layer that accepted the temporal sequences, multiple hidden layers with memory cells to capture intricate patterns over time, and an output layer tasked with predicting the next day's open price. The training process incorporated loss computation at each stage to optimize the model's performance iteratively. Techniques like dropout regularization and L2 regularization were applied to mitigate overfitting, while an early stopping function ensured that training was terminated once validation performance ceased to improve, preventing unnecessary computations and reducing the risk of overfitting.

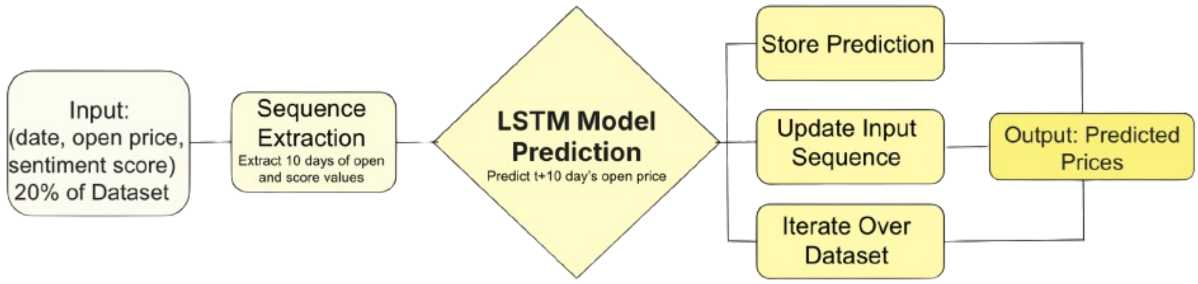

Figure 2: testing process.

The Testing process as seen from Figure 2 involved using the remaining 20% of the dataset, where input sequences of sentiment scores and financial metrics were extracted for evaluation. The LSTM model predicted the open price for the next 10-day period iteratively by updating the input sequence and storing predictions at each step. This iterative testing process validated the model's ability to handle real-world scenarios effectively.

Additionally, hyperparameter tuning was conducted to optimize learning rate, batch size, and the number of LSTM units, maximizing the model's predictive accuracy. This robust training pipeline, supported by preprocessing, regularization, and hyperparameter optimization, enabled the model to capture and analyze intricate temporal patterns. The combination of historical market data and sentiment trends made the model particularly adept at forecasting Bitcoin price movements, delivering actionable insights in a highly volatile and dynamic financial landscape.

4 Results and Discussion

4.1 Experimentation protocol/Evaluation

The loss values for both training and validation datasets are computed using a loss function, typically the **Mean Squared Error (MSE)**, defined as:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2, \quad \text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$

where N is the number of samples, y_i is the actual value for the i-th sample, and $\hat{y}_i$ is the predicted value for the i-th sample. In this case, the actual price and the predicted prices are computed using the sentiment score as a feature along with the open/close price.

$$\text{Training Loss} = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} \mathcal{L}(y_i, \hat{y}_i), \quad \text{Validation Loss} = \frac{1}{N_{\text{val}}} \sum_{i=1}^{N_{\text{val}}} \mathcal{L}(y_i, \hat{y}_i)$$

where N_{train} and N_{val} are the number of training and validation samples, respectively, and $\mathcal{L}(y_i, \hat{y}_i)$ is the loss function (e.g., MSE or MAE).

$$\text{Batch Loss}_k = \frac{1}{|B_k|} \sum_{i \in B_k} \mathcal{L}(y_i, \hat{y}_i), \quad \text{Training Loss (epoch)} = \frac{1}{K} \sum_{k=1}^{K} \text{Batch Loss}_k$$

where K is the total number of batches in an epoch and $|B_k|$ is the number of samples in batch B_k. In our analysis, in order to mitigate a lengthy runtime, we fluctuated the batch sizes from 32 and 64.

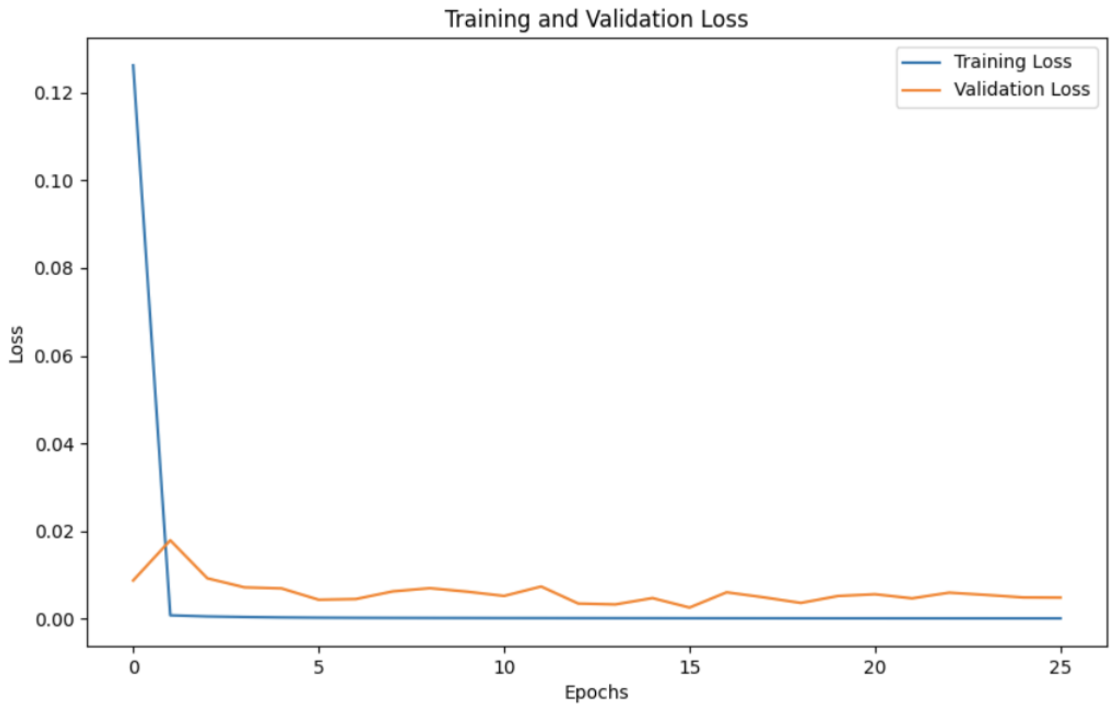

Figure 3: Predicted Prices vs Real Prices (Daily Average).

Figure 3 illustrates the loss values for both training and validation datasets across epochs during the model's training process. The training loss, represented by the blue line, decreases sharply in the initial epochs, indicating rapid learning by the model. After this steep decline, the training loss stabilizes at near-zero levels, reflecting the model's ability to effectively minimize errors on the training dataset. Similarly, the validation loss, shown as the orange line, begins at a higher value but also decreases initially before stabilizing with minimal fluctuations. This behavior suggests that the model generalizes well to unseen data and avoids overfitting, likely due to the application of the L2 regularization technique as well as the dropout and early stopping function.

Further testing was conducted with sentiment scores derived from VADER, BERT, and TextBlob to assess the LSTM model's ability to integrate different sentiment analysis tools into its predictions. Across these experiments, training and validation loss trends exhibited a similar trajectory, reinforcing the model's consistent capability to leverage sentiment scores effectively. Among these, BERT-based sentiment scores achieved the lowest validation loss, highlighting its strength in capturing complex and nuanced sentiment patterns. While TextBlob and VADER also performed well, their simpler methodologies resulted in slightly higher validation losses in some cases.

Overall, the stabilization of both training and validation losses across all tests underscores the robustness of the LSTM model and its ability to incorporate sentiment-driven features for accurate Bitcoin price predictions. These findings validate the utility of combining advanced sentiment analysis tools with financial modeling to enhance prediction performance for such volatile markets.

4.2 Data tables

Date	Actual Price	Predicted Price (TextBlob)	Predicted Price (BERT)	Predicted Price (VADER)	Predicted Price (Random Forest)
2023-05-02	28,262.82347	27,981.168	27,731.752	27,643.627	27,970.684
2023-05-25	26,302.60422	26,071.861	26,137.412	26,133.568	26,302.148
2023-06-04	27,161.53055	26,916.408	26,887.887	26,866.453	27,081.662

Figure 4: Table Of Comparison of Actual and Predicted Prices Using Various Models On Randomly Selected Days.

Sentiment Tool	Accuracy
TextBlob	56%
VADER	33%
BERT	45%
Random Forest (weighted combination of TextBlob, VADER, and BERT scores)	78%

Figure 5: Table Of Accuracy of Each Sentiment Tool Used.

The table, in Figure 4, compares the actual price of a financial asset with the predicted prices from four different sentiment tools: TextBlob, BERT, VADER, and Random Forest, across three different randomly selected dates in May and June 2023. On all three dates, the predicted prices from the models closely align with the actual prices, with some minor discrepancies. For example, on May 2, 2023, the actual price was 28,262.82, and the predictions ranged from 27,643.63 (VADER) to 27,981.17 (TextBlob), showing small differences.

However, the slight variations in predictions across the models suggest that some models may perform better under different conditions or with different types of data. For instance, the Random Forest model showed a slight overestimation on May 2 and June 4, while VADER slightly underpredicted on the same dates. This variance could point to the strengths and weaknesses of each model, with Random Forest potentially benefiting from more data-driven feature analysis, while simpler models like TextBlob may not capture all underlying market dynamics. Overall, all models seem reasonably accurate, though a more extensive dataset and testing might reveal further nuances in their predictive capabilities.

Figure 5 shows the accuracy results highlighting the strengths and weaknesses of individual sentiment analysis tools compared to an ensemble approach. TextBlob, VADER, and BERT achieve accuracies of 56%, 33%, and 45%, respectively, indicating limited performance when used independently. Notably, TextBlob outperforms the others, but its accuracy is still relatively low.

The ensemble model, which uses a Random Forest algorithm to weight and combine the scores of TextBlob, VADER, and BERT, achieves a significantly higher accuracy of 78%. This improvement demonstrates the effectiveness of ensemble methods in leveraging the strengths of multiple models while mitigating their individual weaknesses. It suggests that integrating diverse sentiment analysis approaches can provide a more reliable and accurate sentiment prediction.

4.3 Graphs

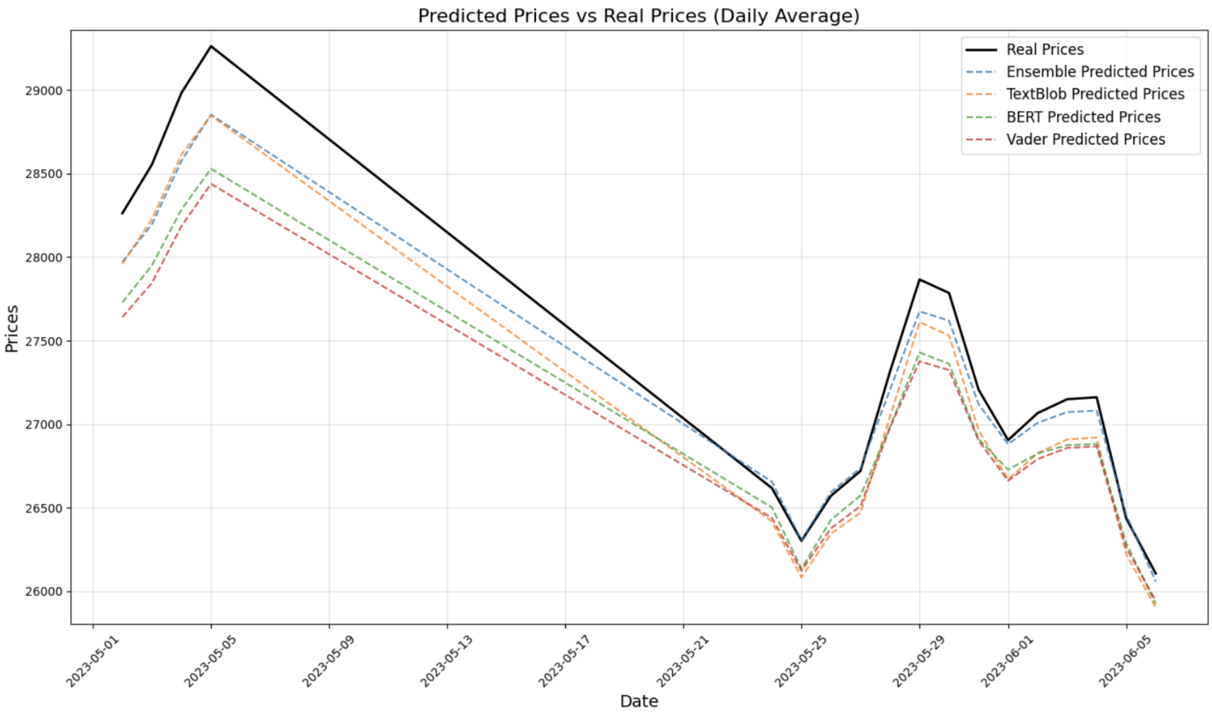

Figure 6: Predicted Prices vs Real Prices (Daily Average).

The comparative analysis of actual Bitcoin prices against predicted prices generated by various models over a specific time frame. This can be seen in Figure 6, which visualizes the effectiveness of sentiment-based prediction models and an ensemble approach in capturing market trends.

The real prices, represented by a solid black line, exhibit significant fluctuations over the observed period. These fluctuations include sharp increases and decreases, characteristic of Bitcoin's well-documented volatility. The dynamic nature of these price movements sets a challenging benchmark for prediction models, making accurate forecasting a complex task.

The graph includes predictions from several models:

- The **Random Forest ensemble model**, represented by the blue dashed line, closely tracks the real prices. This suggests that combining predictions from multiple models enhances the robustness and accuracy of the forecasts.

- The **TextBlob model**, depicted by the orange dashed line, provides moderately accurate predictions but shows deviations from the real prices in certain areas, particularly during periods of rapid price movement.

- The **BERT model**, illustrated by the green dashed line, demonstrates strong performance, with predictions that stay relatively close to the trends of the real prices.
- The **VADER model**, shown as the red dashed line, also captures the general direction of price changes but exhibits slightly larger deviations compared to the ensemble and BERT models.

All models demonstrate the ability to capture the general trends in Bitcoin price movements, including peaks and troughs. However, notable discrepancies are observed during periods of rapid price changes, where the predicted prices tend to lag slightly behind the real prices. This lag highlights the inherent challenges in modeling sudden and volatile market shifts.

Overall, the ensemble model outperforms the individual sentiment-based models (TextBlob, BERT, and VADER) in terms of accuracy and reliability. This comparison underscores the advantages of integrating diverse predictive techniques into an ensemble framework, enabling better alignment with real market dynamics while addressing the limitations of single-model predictions.

5 Discussion

The findings from our analysis demonstrate the potential of integrating sentiment analysis with financial metrics to predict Bitcoin price movements. Bitcoin's extreme price volatility and the lack of consideration for sentiment factors in traditional financial models provided a significant motivation for this research. By addressing these gaps, our study contributes to the limited body of research linking social media sentiment with predictive financial modeling.

Our methodology involved processing 285,000 tweets collected between January 31, 2023, and June 6, 2023, using three sentiment analysis tools—VADER, TextBlob, and BERT. These tools allowed us to derive sentiment scores that encapsulated public opinions about Bitcoin, classified into positive, neutral, and negative sentiments. These sentiment indicators were then merged with Bitcoin price data, ensuring alignment by date to create a unified dataset that combined social sentiment and financial market factors.

The LSTM model, tailored for time-series forecasting, was trained using this dataset. A robust training pipeline included preprocessing steps such as sequence extraction, normalization, and the use of a sequence length of 10 days for input. Both training and validation loss graphs indicated strong model performance, with minimal overfitting achieved through dropout regularization and early stopping.

Testing revealed that the ensemble model, which integrated predictions from all three sentiment analysis tools, consistently outperformed individual models. While all models captured general price trends effectively, the BERT-based sentiment scores showed slightly better validation loss, highlighting their ability to capture nuanced patterns in sentiment data. VADER and TextBlob models, while slightly less accurate, still demonstrated the utility of simpler sentiment analysis techniques for price prediction tasks.

Our findings highlight that sentiment-driven features, when combined with financial metrics, can significantly improve the accuracy of predictive models in volatile markets like Bitcoin. The results emphasize the importance of leveraging diverse sentiment analysis tools and ensemble techniques to create robust predictive frameworks. Future work could expand on this approach by incorporating additional sentiment data sources or refining the models to address rapid price shifts more effectively.

6 Conclusion

This study demonstrates the potential of integrating sentiment analysis with financial metrics to predict Bitcoin price movements. By combining the strengths of multiple sentiment analysis tools—VADER, TextBlob, and BERT—along with key financial indicators such as price fluctuations and trading volume, we developed a robust ensemble model capable of forecasting Bitcoin's volatile price behavior with improved accuracy. Our model significantly outperformed individual sentiment analysis models, achieving an accuracy of 78%, showcasing the effectiveness of ensemble methods in mitigating the weaknesses of single models.

The use of temporal sequences and an LSTM network further enhanced the model's ability to capture time-dependent patterns in market behavior, enabling it to make accurate predictions based on past sentiment and price trends. This research contributes to the growing body of work in predictive financial modeling by highlighting the value of incorporating social media sentiment in cryptocurrency forecasting.

Despite the promising results, this study has several limitations. First, the sentiment analysis tools, while effective, may not fully capture the nuances of financial market sentiment, especially when dealing with ambiguous or contradictory language in social media posts. Additionally, the model's reliance on historical price data means that it may struggle to adapt to sudden, unforeseen market events or structural shifts in the cryptocurrency market. Moreover, while the ensemble model improved accuracy, it still remains limited by the quality of the input data and the biases inherent in social media sentiment.

Future work could extend this approach by integrating additional sentiment data sources, such as news articles or other social media platforms, to refine the model's predictive capabilities. Incorporating more granular data, such as user demographics or market sentiment on specific subreddits or Twitter accounts, may provide more targeted insights. Furthermore, exploring more advanced techniques in natural language processing and machine learning, such as transformer-based models or reinforcement learning, could lead to even more accurate predictions, especially during periods of extreme market volatility. Additionally, it would be valuable to test the model on other cryptocurrencies to assess its generalizability and robustness across different digital assets.

This study serves as a stepping stone towards more reliable, data-driven approaches for understanding and predicting cryptocurrency markets, with the potential for further refinement to improve both accuracy and adaptability in real-world applications.

References

[1] O. Sattarov, H. S. Jeon, R. Oh, and J. D. Lee, "Forecasting Bitcoin Price Fluctuation by Twitter Sentiment Analysis," 2020 International Conference on Information Science and Communications Technologies (ICISCT), Nov. 2020, https://doi.org/10.1109/icisct50599.2020.9351527.

[2] J. V. Critien, A. Gatt, and J. Ellul, "Bitcoin price change and trend prediction through twitter sentiment and data volume," Financial Innovation, vol. 8, no. 1, May 2022, https://doi.org/10.1186/s40854-022-00352-7.

[3] I. Georgoula, D. Pournarakis, C. Bilanakos, D. N. Sotiropoulos, and G. M. Giaglis, "Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices," SSRN Electronic Journal, 2015, https://doi.org/10.2139/ssrn.2607167.

[4] I. Gurrib and F. Kamalov, "Predicting bitcoin price movements using sentiment analysis: a machine learning approach," Studies in Economics and Finance, vol. ahead-of-print, no. ahead-of-print, Dec. 2021, https://doi.org/10.1108/sef-07-2021-0293.

[5] N. Passalis et al., "Multisource financial sentiment analysis for detecting Bitcoin price change indications using deep learning," Neural Computing and Applications, Jul. 2022, https://doi.org/10.1007/s00521-022-07509-6.

[6] D. R. Pant, P. Neupane, A. Poudel, A. K. Pokhrel, and B. K. Lama, "Recurrent Neural Network Based Bitcoin Price Prediction by Twitter Sentiment Analysis," 2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS), Oct. 2018, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8586824.

[7] G. Serafini et al. "Sentiment-Driven Price Prediction of the Bitcoin based on Statistical and Deep Learning Approaches," IEEE Xplore, Jul. 01, 2020, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9206704.

[8] A. Bâra and S.-V. Oprea, "An ensemble learning method for Bitcoin price prediction based on volatility indicators and trend," Engineering Applications of Artificial Intelligence, vol. 133, pp. 107991–107991, Jul. 2024, https://doi.org/10.1016/j.engappai.2024.107991.

[9] A. M. Rather, "A new method of ensemble learning: case of cryptocurrency price prediction," Knowledge and Information Systems, Dec. 2022, https://doi.org/10.1007/s10115-022-01796-0.

[10] E. Sin and L. Wang, "Bitcoin price prediction using ensembles of neural networks," 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, China, 2017, pp. 666-671, 10.1109/FSKD.2017.8393351.

[11] Z. Ye, Y. Wu, H. Chen, Y. Pan, and Q. Jiang, "A Stacking Ensemble Deep Learning Model for Bitcoin Price Prediction Using Twitter Comments on Bitcoin," Mathematics, vol. 10, no. 8, p. 1307, Apr. 2022, https://doi.org/10.3390/math10081307.

[12] M. Shin, D. Mohaisen, and J. Kim, "Bitcoin Price Forecasting via Ensemble-based LSTM Deep Learning Networks," 2021 International Conference on Information Networking (ICOIN), Jan. 2021, https://doi.org/10.1109/icoin50884.2021.9333853.

[13] I. E. Livieris, E. Pintelas, S. Stavroyiannis, and P. Pintelas, "Ensemble Deep Learning Models for Forecasting Cryptocurrency Time-Series," Algorithms, vol. 13, no. 5, p. 121, May 2020, https://doi.org/10.3390/a13050121.

[14] R. G. da Silva, M. H. Dal Molin Ribeiro, N. Fraccanabbia, V. C. Mariani, and L. dos Santos Coelho, “Multi-step ahead Bitcoin Price Forecasting Based on VMD and Ensemble Learning Methods,” IEEE Xplore, Jul. 2020, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9207152.

[15] H. Htun, M. Biehl, and N. Petkov, “Survey of feature selection and extraction techniques for stock market prediction,” Financial Innovation, vol. 9, no. 1, Jan. 2023, https://doi.org/10.1186/s40854-022-00441-7.

[16] S. Singh, A. Pise, and B. Yoon, “Prediction of bitcoin stock price using feature subset optimization,” Heliyon, vol. 10, no. 7, p. e28415, Apr. 2024, https://doi.org/10.1016/j.heliyon.2024.e28415.

[17] Li, Yuanhang & Xie, Zhengjie, “A Study on CNN Feature Extraction for Stock Price Prediction”. BCP Business & Management. 33. 326-330, 2022, 10.54691/bcpbm.v33i.2771.

[18] Jagadesh, B.N., RajaSekhar Reddy, N.V., Udayaraju, P. et al. “Enhanced stock market forecasting using dandelion optimization-driven 3D-CNN-GRU classification”. Sci Rep 14, 20908, 2024, https://doi.org/10.1038/s41598-024-71873-7.

[19] J. Wang, J. He, C. Feng, L. Feng, and Y. Li, “Stock index prediction and uncertainty analysis using multi-scale nonlinear ensemble paradigm of optimal feature extraction, two-stage deep learning and Gaussian process regression,” Applied Soft Computing, vol. 113, p. 107898, Dec. 2021, https://doi.org/10.1016/j.asoc.2021.107898.

[20] J. Chen, “Analysis of Bitcoin Price Prediction Using Machine Learning,” Journal of Risk and Financial Management, vol. 16, no. 1, p. 51, Jan. 2023, https://www.mdpi.com/1911-8074/16/1/51.

[21] R. Gupta and J. E. Nalavade, “Metaheuristic Assisted Hybrid Classifier for Bitcoin Price Prediction,” Cybernetics and Systems, pp. 1–25, Oct. 2022, 10.1080/01969722.2022.2129376.