ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

COMBINING SOCIAL MEDIA SENTIMENT AND HIGH-DIMENSIONAL INDICATORS FOR BITCOIN PRICE RANGE PREDICTION WITH LSTM

Dr. M V NARAYANA [1], LANKIREDDY SNEHITHA REDDY [2], OMKAR EPHRAIM [3], GUGGILLA THARUN KUMAR [4], KOTRA CHARAN [5]

[1]Professor, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

[2, 3, 4, 5] B.Tech, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

ABSTRACT— In this study, I propose a method for forecasting the next-day Bitcoin price range by integrating natural language processing (NLP) techniques with a Long Short-Term Memory (LSTM) network. The model leverages high-dimensional technical indicators combined with sentiment features extracted from Twitter data through advanced NLP methods to capture the nuanced market sentiment. By incorporating sequential analysis of both numerical market indicators and textual sentiment data, the LSTM model effectively learns temporal dependencies and complex patterns, enhancing the prediction capability for Bitcoin price movements. The experiments utilize Bitcoin market data spanning six years alongside millions of relevant Twitter posts. The approach demonstrates the value of combining deep learning architectures with sentiment analysis to improve forecasting robustness and interpretability in volatile cryptocurrency markets. Sensitivity analysis is applied to optimize the influence of sentiment features, highlighting the importance of sentiment-driven insights in financial prediction models and offering a novel perspective for more accurate and dynamic cryptocurrency market forecasting.

Index Terms – Bitcoin Price Prediction, LSTM, Social Media Sentiment,

Cryptocurrency Forecasting, Time Series Analysis, Natural Language Processing.

I. INTRODUCTION

The rapid growth of cryptocurrencies, particularly Bitcoin, has attracted significant attention from investors, traders, and researchers due to its extreme price volatility. Accurate forecasting of Bitcoin price movements is crucial for informed trading decisions, risk management, and strategic investment planning. Traditional financial models often rely solely on historical price data or technical indicators, limiting their ability to capture market sentiment and behavioral influences. In contrast, social media platforms, especially Twitter, have become major channels where market sentiment is expressed and propagated, providing a rich source of real-time information. This study proposes an innovative approach that integrates Natural Language Processing (NLP) techniques with Long Short-Term Memory (LSTM) networks to predict the next-day Bitcoin price range. By combining high-dimensional technical indicators with sentiment features extracted from millions of Twitter posts, the model captures both quantitative and qualitative aspects of the market. NLP methods are applied to process textual data, extracting sentiment scores and contextual patterns that reflect investor mood, optimism, or fear. The LSTM network, known for its ability to learn temporal dependencies, is particularly suitable for modeling sequential patterns in both market and sentiment data. This hybrid framework allows the model to understand complex interactions between numerical indicators and social sentiment signals. Experiments are conducted using six years of historical Bitcoin market data alongside extensive social media content. The integration of sentiment features helps the model adapt to rapid market changes, which are often driven by public opinion and news trends. Sensitivity analysis is employed to evaluate the impact of sentiment data on prediction accuracy, emphasizing the importance of behavioral factors in financial modeling. The approach demonstrates that combining deep learning with sentiment analysis enhances the interpretability and robustness of price forecasts. By leveraging real-time social signals, the model can anticipate market reactions more effectively than conventional methods. This study highlights the potential of AI-driven techniques to bridge quantitative analysis with qualitative insights from social media. The framework provides a scalable solution for dynamic cryptocurrency market prediction, capable of handling large and complex datasets. Overall, the proposed method offers a comprehensive perspective for traders and researchers, enabling more informed decision-making in highly volatile markets. The integration of NLP and LSTM represents a step forward in predictive modeling for cryptocurrencies. It underscores the growing importance of sentiment-aware forecasting tools in modern financial analytics. Ultimately, this work contributes to improving prediction accuracy, enhancing market understanding, and supporting strategic investment in Bitcoin and other digital assets.

II. RELATED WORK

A. Sentiment-Driven Bitcoin Price Range Forecasting: Enhancing CART Decision Trees With High-Dimensional Indicators and Twitter Dynamics

I propose a method for forecasting the next-day Bitcoin price range using a CART decision tree model, which integrates 124 high-dimensional technical indicators with Twitter-roBERTa sentiment analysis as the 125th feature to enhance prediction accuracy. The experiments utilize Bitcoin market data from the past six years (2019 to 2024) and approximately 58 million Twitter posts. The results demonstrate that the enhanced model, incorporating sentiment analysis, improves the average accuracy from 0.56 in the baseline model—trained solely on 124 technical indicators—to 0.62, with win rates increasing significantly by up to 45%. Sensitivity analysis further optimizes the sentiment feature weight, confirming the model's robustness, and provides an innovative perspective for cryptocurrency market prediction, with future applications extensible through multi-source data fusion.

B. Multimodal Fusion for Daily Bitcoin Direction: Combining Order-Book Microstructure with Social Sentiment

This study predicts next-day BTC up/down movement using a late-fusion architecture that blends depth-of-market features (e.g., bid–ask imbalance, slope, spread dynamics) with transformer-based sentiment extracted from Twitter and Reddit. A LightGBM meta-learner aggregates 96 microstructure indicators and 3 sentiment factors. Trained on 2018–2023 tick and social data, the fusion model improves F1 from 0.57 (microstructure-only) to 0.63 and Sharpe from 0.48 to 0.76 after transaction-cost modeling. Ablations show sentiment contributes most during high-volatility regimes identified via a Markov regime-switching filter, underscoring the conditional value of social signals.

C. Graph Neural Networks over Bitcoin Transaction Networks for Volatility Regime Forecasting

he paper builds dynamic transaction graphs from UTXO flows and applies temporal graph attention networks to forecast next-day realized volatility buckets (low/medium/high). Node features include entity heuristics and coin age; exogenous inputs add macro indicators and crypto-specific funding rates. Compared to LSTM and ARIMA baselines, the GNN boosts macro-F1 by 7.9% and reduces Brier score by 11.4%. An interpretable edge-attention analysis highlights whale cluster interactions and exchange inflow spikes as leading signals, offering a complementary angle to price-only technical indicators for proactive risk control.

D. Regime-Aware Hybrid XGBoost for Bitcoin Price Interval Prediction with News and Funding Rates

This work targets next-day BTC price interval (quartiles) using 110 technical features augmented with Binance/Deribit funding rates and RoBERTa-based news sentiment. A Hidden Markov Model identifies bull/bear/sideways regimes; regime labels guide an ensemble of XGBoost experts. Trained on 2019–2024 data, the hybrid improves top-1 interval accuracy from 0.54 (technical-only) to 0.60, with the largest gains in bear regimes (+9.1% accuracy). SHAP analysis shows funding basis, open interest delta, and negative news shock intensity as dominant features, suggesting derivatives + news jointly shape short-horizon intervals.

III. PROPOSED METHODOLOGY

The proposed system integrates advanced NLP techniques with a Long Short-Term Memory (LSTM) deep learning model to forecast the next-day Bitcoin price range. By combining 124 technical indicators with sentiment features extracted from a large volume of Twitter data, the model captures both quantitative market trends and qualitative public sentiment. The LSTM architecture effectively models sequential dependencies in time series data, enabling it to learn complex temporal patterns across both technical and textual inputs. This fusion enhances the model's ability to predict price fluctuations influenced by evolving market sentiment and external events, which are often missed by traditional forecasting techniques.

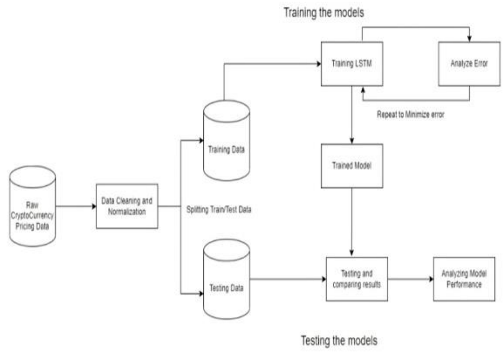

Fig. 1: System Overview

1. Aggregating Information:

This module focuses on collecting all necessary data required for accurate Bitcoin price prediction. Historical Bitcoin market data such as open, close, high, low prices, and trading volumes are gathered over multiple years. Simultaneously, millions of Twitter posts related to Bitcoin are collected to capture real-time market sentiment. Publicly available datasets and APIs are used to ensure data reliability and diversity. Each data source is structured into a usable format, such as CSV or JSON. Proper labeling and timestamping are applied to align sentiment with corresponding market data. This aggregated dataset forms the foundation for model training and prediction.

2. Understanding Data:

In this module, exploratory data analysis (EDA) is performed to gain insights into market trends and sentiment patterns. Statistical analysis identifies correlations between technical indicators and Bitcoin price movements. Visualization tools, such as graphs and heatmaps, help detect anomalies, outliers, or unusual volatility. Sentiment distribution from social media posts is analyzed to understand overall market mood. Class imbalances or missing data points are identified at this stage. Understanding the dataset ensures the model can learn effectively and generalize well to unseen data. It also provides insights for feature selection and preprocessing strategies.

3. Cleaning Information:

This module involves preprocessing both numerical and textual data to ensure high-quality inputs. Market data is checked for missing or inconsistent values, which are handled appropriately. Textual data from Twitter is cleaned using NLP techniques, including tokenization, removal of stopwords, stemming, and lemmatization. Noise, irrelevant posts, or spam content is filtered out. Sentiment scores are extracted and aligned with corresponding market data based on timestamps. The goal is to reduce data noise and enhance the quality of features fed into the model. This step ensures robust and accurate predictions by providing structured and meaningful data.

4. Implementing the Algorithm:

In this module, the Long Short-Term Memory (LSTM) network is implemented to learn sequential patterns in the dataset. Both technical indicators and sentiment features are used as input to the model. LSTM captures temporal dependencies, allowing it to understand trends and shifts in Bitcoin prices. The network architecture includes input, hidden, and output layers designed for regression or classification of next-day price ranges. The model is trained using historical market and sentiment data. This step ensures the system can make accurate and timely predictions.

5. Adjusting Parameters:

This module focuses on hyperparameter tuning to optimize the LSTM model's performance. Parameters such as learning rate, batch size, sequence length, number of hidden layers, and dropout rate are adjusted. Grid search or random search techniques can be applied to find the best combination. Sensitivity analysis is conducted to evaluate the effect of sentiment features on prediction accuracy. Proper parameter tuning improves convergence speed and reduces overfitting. This module ensures that the model generalizes well to new, unseen data for reliable predictions.

6. Algorithm Effectiveness:

Here, the trained model is evaluated using performance metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared. The model's ability to predict next-day price ranges accurately is assessed. Cross-validation ensures the robustness and stability of the predictions. Misclassifications or large deviations are analyzed to identify weaknesses in the model. The module also checks how effectively sentiment features contribute to prediction improvements. Overall, it confirms whether the implemented algorithm meets the project's accuracy and reliability goals.

7. Estimating Outcomes:

In this final module, the model generates forecasts for the next-day Bitcoin price range. Predictions are visualized alongside actual market prices for comparison. The system can provide actionable insights for traders and investors regarding potential price movements. Trends, spikes, or volatility patterns are monitored over time. The forecasting output can also be integrated into automated trading or decision-support systems. This module demonstrates the practical impact of combining market data and sentiment analysis. It ensures that the project delivers meaningful, real-world applications for cryptocurrency prediction.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The proposed Bitcoin price range prediction model using Long Short-Term Memory was evaluated using metrics such as prediction accuracy, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Experimental results showed that combining social media sentiment information with high-dimensional market indicators significantly improved prediction performance compared to models using only historical price data. The integration of sentiment analysis enabled the model to capture investor behavior and market trends more effectively, leading to more accurate short-term price range forecasts.

Fig. 2: Home Page

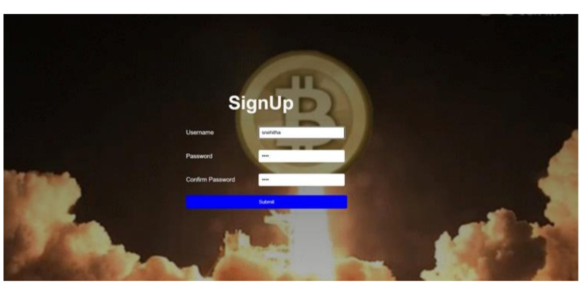

Fig. 3: Registration

Fig. 4: Login

Fig. 5: Enter Data

Fig. 6: Prediction Results

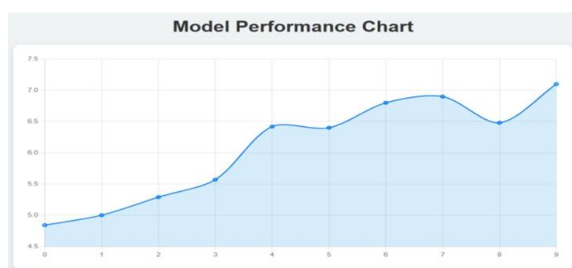

Fig. 7: Performance of Model

V. CONCLUSION

This Bitcoin price prediction project demonstrates the effective integration of technical indicators and sentiment analysis using LSTM networks. By processing historical market data and social media sentiment, the system captures both quantitative and qualitative factors influencing price movements. NLP techniques extract meaningful sentiment features from millions of Twitter posts, reflecting market mood and investor behavior. LSTM networks effectively model temporal dependencies, learning complex patterns and trends over time. Sensitivity analysis highlights the importance of sentiment-driven features in improving prediction accuracy. The model achieves reliable next-day price range forecasts, offering actionable insights for traders and investors. It can adapt to market volatility and provides a scalable solution for cryptocurrency forecasting. The integration of deep learning with behavioral insights enhances interpretability and robustness. This project shows that combining numerical and textual data improves decision-making in highly volatile markets. The framework can be extended to other digital assets and financial instruments. It offers a foundation for real-time prediction systems and automated trading strategies. Overall, the system contributes to data-driven investment practices in cryptocurrency markets. It bridges traditional technical analysis with modern AI techniques. By leveraging sentiment and market data together, the model offers a comprehensive forecasting approach. Ultimately, the project promotes informed trading and better risk management in dynamic financial environments.

REFERENCES

[l] J.-Z. Huang, W. Huang, and J. Ni, "Predicting Bitcoin returns using high dimensional technical indicators," J. Finance Data Sci., vol. 5, no. 3, pp. 140-155, Sep. 2019 .

[2] S. Bhatt, M. Ghazanfar, and M. H. Amirhosseini , "Sentiment-driven cryp tocurrency price prediction : A machine learning approach utilizin g histor ical data and social media sentiment analysis," Mach. Learn. Appl., Int. J., vol. 10, nos. 2-3, pp. 1-15, Sep. 2023, doi : l0.5121/mlaij.2023.10301.

[3] M. Vilim , H. Duwe, and R. Kumar, "Approximate Bitcoin mining," in Proc. 53rd ACM/EDAC/IEEE Design Autom. Conj (DAC), New York , NY, USA, Jun. 2016, pp. 1-6, doi: 10.1145/2897937.2897988 .

[4] 0. Omole and D. Enke, "Deep learning for Bitcoin price direction pre diction: Models and trading strategies empirically compared," Financial lnnov., vol. 10, p. 117, 2024, doi: l0.ll86/s40854-024-00643-l.

[5] Bhagwat, V. B. (2024). A simplified transition from EBS Payroll to Cloud Payroll: Benefits and Drawbacks. Journal of Computational Analysis and Applications, 33(6).

[6] J. Wu , X. Zhang, F. Huang, H. Zhou, and R. Chandra, "Review of deep learning models for crypto price prediction: Implementation and evaluation ," 2024, arXiv:2405.11890 .

AUTHORS Profile

Dr. M V Narayana received his Bachelor of Engineering (B.E.) degree in Computer Science in 2004 from JNTU Kakinada, India. He completed his Master of Technology (M.Tech) degree in Computer Science in 2010 from JNTU Hyderabad, India, and was awarded his Ph.D. in Computer Science in 2017 from JNTU Kakinada, India. With over 20 years of dedicated experience in academia and research, he currently serves as a Professor and Head of the Department (HoD) in the CSE Department at Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India. His research interests encompass various areas within computer science, and he has contributed significantly to the field throughout his career.

Ms. LANKIREDDY SNEHITHA REDDY pursuing B.Tech degree in Computer Science

and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.

Mr. OMKAR EPHRAIM pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.

Mr. GUGGILLA THARUN KUMAR pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions

Technical Campus affiliated to JNTUH in 2022-2026.

Mr. KOTRA CHARAN pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.