JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A HYBRID CRYPTO TREND ANALYSIS AND ONLINE GRAPHICAL REPRESENTATION MODEL USING NEURAL NETWORK

[1]Dr. S. Rajesh, [2]Keerthivasan V, [3]Madhesh V, [4]Raghul B
[1]Assistant Professor, [2,3,4]Student
[1,2,3,4]Department of Computer Science and Engineering,
[1,2,3,4]Sri Ramakrishna Institute of Technology, Coimbatore, India

Abstract : Bitcoin price prediction has become a significant area of research due to its volatile nature and potential for financial gain. In this study, we employed four distinct algorithms, namely Auto Regressive Integrated Moving Average with Exogenous Variables (ARIMAX), Long Short-Term Memory (LSTM) networks, Facebook Prophet, and XGBoost, to forecast the price of Bitcoin. The analysis utilized historical Bitcoin price data spanning several years. Initially, the data underwent preprocessing steps, including handling missing values and visualizing temporal trends. Subsequently, we engineered features such as rolling means and standard deviations to capture potential patterns. Each algorithm was then applied, and their respective predictions were evaluated using performance metrics like Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Our results indicate that LSTM and FB Prophet demonstrated superior predictive capabilities, yielding the highest accuracy percentages. Additionally, we explored the potential of hybrid models by combining forecasts from multiple algorithms, which further enhanced prediction accuracy. Overall, our study contributes to the ongoing efforts to develop robust models for Bitcoin price prediction, thereby aiding investors and stakeholders in making informed decisions in the cryptocurrency market.

***IndexTerms* - Cryptocurrency, price analysis, volume analysis, deep learning, neural network, prediction methods, hybrid model.**

I.INTRODUCTION

Bitcoin, the pioneering cryptocurrency introduced by an unknown person or group of people using the pseudonym Satoshi Nakamoto in 2009, [1] has garnered immense attention and investment interest over the past decade. Operating on a decentralized peer-to-peer network, Bitcoin offers a revolutionary alternative to traditional financial systems, enabling secure, transparent, and pseudonymous transactions without the need for intermediaries like banks or governments. Its decentralized nature, powered by blockchain technology [2], ensures immutability and resistance to censorship, making it an attractive asset for investors seeking diversification and potential high returns.As the cryptocurrency market continues to evolve and mature, the need for accurate price forecasting tools becomes increasingly crucial.Predicting the price movements of Bitcoin presents a complex challenge due to its inherent volatility [24], influenced by a myriad of factors such as market sentiment, regulatory developments, technological advancements, and macroeconomic trends. To address this challenge, various methodologies and algorithms have been proposed to forecast Bitcoin prices, ranging from traditional time series analysis techniques to advanced machine learning models.In this study, we employ a comprehensive approach that integrates both classical time series analysis and cutting-edge machine learning algorithms to forecast the price of Bitcoin. Specifically, we leverage the Auto Regressive Integrated Moving Average with Exogenous Variables (ARIMAX) model [6], which is widely used in financial forecasting due to its ability to capture both linear and non-linear relationships in time series data. Additionally, we harness the predictive power of Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN) known for its ability to capture long-term dependencies in sequential data, making it well-suited for modeling complex time series like Bitcoin prices[5].Furthermore, we incorporate the Facebook Prophet algorithm[19], developed by Facebook's Core Data Science team, which is specifically designed for forecasting time series data with strong seasonal patterns and irregularities. Lastly, we utilize XGBoost [18], an ensemble learning algorithm based on decision trees, known for its efficiency and effectiveness in handling structured data and making accurate predictions.By combining these diverse methodologies, our study aims to provide a robust and accurate forecasting framework for Bitcoin prices, contributing to the ongoing efforts to understand and predict the dynamics of the cryptocurrency market. Through rigorous evaluation and comparison of the performance of each algorithm, we seek to identify the most effective approaches for Bitcoin price prediction, ultimately empowering investors and stakeholders with valuable insights for decision-making in the volatile cryptocurrency landscape.

II. SYSTEM ARCHITECTURE

The system architecture for the hybrid model comprises several key components. Initially, raw time-series data is collected and preprocessed to handle missing values and outliers. Feature engineering is then performed to extract relevant attributes from the preprocessed data.Four distinct models, ARIMAX, FB Prophet, XG BOOST and LSTM are individually trained using historical data to capture linear trends and nonlinear patterns, respectively. Following training, these models are integrated into a hybrid architecture[9], potentially using techniques like stacking or averaging to combine their predictions.The hybrid model's performance is evaluated using validation data, employing metrics such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE) [20]. Upon satisfactory evaluation, the hybrid model is deployed into production, where it generates real-time forecasts. Monitoring tools are implemented to track model performance and data anomalies, while retraining ensures the model remains accurate and up-to-date with the latest trends.

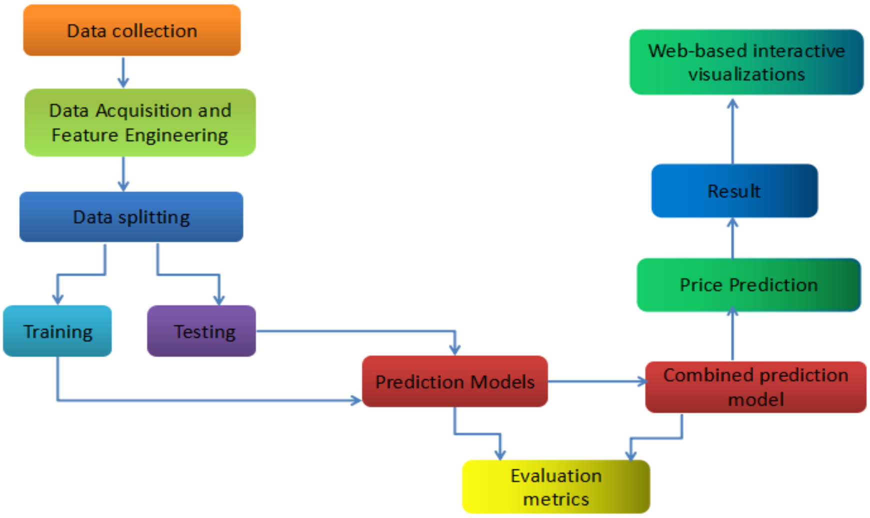

Figure 1. System Architecture.

III. METHODOLOGY

Data Collection and Preprocessing: We begin by collecting historical Bitcoin price data from reputable cryptocurrency exchanges, ensuring a comprehensive dataset spanning multiple years to capture diverse market conditions and trends[23].The collected data includes daily opening, closing, high, and low prices, as well as trading volume and other relevant indicators. To ensure data accuracy and reliability, we meticulously clean and preprocess the dataset, handling missing values, outliers, and inconsistencies through techniques such as interpolation,smoothing and normalization[17].

Feature Engineering: In preparation for model training, we conduct feature engineering to extract meaningful input variables that can potentially influence Bitcoin prices. These features may include technical indicators such as moving averages, Rolling Statistics [16], Moving Average Convergence Divergence, Lag plots, ACF, PACF [21] and Bollinger Bands, as well as exogenous features such as rolling mean, standard deviation and moving average. By incorporating a diverse range of features, we aim to capture the multifaceted nature of the cryptocurrency market and improve the predictive power of our models.

Model Selection and Training: We employ a CPD [5] and diverse set of forecasting models, including both traditional statistical methods and modern machine learning algorithms, to explore different approaches for predicting Bitcoin prices. These models include the Auto Regressive Integrated Moving Average (ARIMA) model [6], which is well-suited for capturing linear dependencies and trends in time series data, as well as advanced machine learning techniques such as Long Short-Term Memory (LSTM) networks, which excel at capturing non-linear relationships and long-term dependencies in sequential data. Additionally, we experiment with ensemble learning methods like XGBoost to leverage the collective wisdom of multiple models and enhance prediction accuracy and FB Prophet model well suitable for time series data.

Model Evaluation and Validation: To assess the performance of each forecasting model, we employ rigorous evaluation metrics[20] such as mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE), which provide insights into the accuracy and precision of predictions. We also conduct cross validation experiments to validate the robustness and generalization ability of the models across different time periods and market conditions. By comparing the performance of each model against baseline benchmarks and alternative approaches, we aim to identify the most effective methodologies for Bitcoin price prediction.

Hyperparameter Tuning & Optimization: In order to fine-tune the performance of our forecasting models, we conduct extensive hyperparameter tuning and optimization experiments[22], adjusting model parameters, architecture configurations, and training strategies to maximize prediction accuracy and minimize overfitting. This process involves iteratively experimenting with different parameter settings and evaluating their impact on model performance using techniques such as grid search, random search, and Bayesian optimization. By optimizing model hyperparameters, we seek to enhance the stability and reliability of our predictions, enabling more informed decision-making in cryptocurrency trading and investment strategies.

Through the systematic application of these methodologies, our study aims to advance the state-of-the-art in Bitcoin price forecasting and provide valuable insights for investors, traders, and researchers seeking to navigate the dynamic and volatile cryptocurrency market. By leveraging a diverse range of data sources, features, and modeling techniques, we strive to develop robust and accurate forecasting models capable of capturing the complex dynamics of Bitcoin prices and empowering stakeholders with actionable insights for risk management and portfolio optimization.

ALGORITHM USED

ARIMAX (AutoRegressive Integrated Moving Average with Exogenous Variables): ARIMAX is a time series forecasting model that combines autoregressive (AR) and moving average (MA) components with exogenous variables (X)[6]. In this approach, the past values of the variable being forecasted are regressed on past values of the same variable and past values of other time series that may influence it. The model also incorporates information from external factors or predictors, which can improve forecast accuracy. In the provided code, ARIMAX is applied to predict Bitcoin prices using both historical Bitcoin data and engineered features as exogenous variables.

Prophet: Prophet is a forecasting tool developed by Facebook's Core Data Science team[19]. It is designed to handle time series data with strong seasonal patterns and multiple seasonality. Prophet decomposes time series data into trend,seasonality, and holiday components and models them separately. It can also handle missing data and outliers gracefully. In the provided code, Prophet is utilized to forecast Bitcoin prices based on historical data, with additional features engineered to enhance prediction accuracy.

XGBoost (Extreme Gradient Boosting): XGBoost is an implementation of gradient boosted decision trees designed for speed and performance. It is widely used for supervised learning tasks, including regression and classification, and has gained popularity for its accuracy and efficiency[18]. XGBoost builds a series of decision trees iteratively, each correcting the errors of its predecessor, and combines their predictions to produce a final output.
In the provided code, XGBoost is employed to predict Bitcoin prices using engineered features derived from historical data.

LSTM (Long Short-Term Memory):
LSTM is a type of recurrent neural network (RNN) architecture designed to model sequential data and handle long-term dependencies.Unlike traditional feed forward neural networks,LSTM networks [4] have feedback connections that allow them to process sequences of data over time. They are particularly effective for time series forecasting tasks due to their ability to retain information over long periods. In the provided code, [5] LSTM is used to forecast Bitcoin prices based on historical data, with the model trained to capture complex patterns and dependencies within the time series.

The hybrid model employed in this study combines the predictive capabilities of ARIMAX, XG BOOST, FB PROPHET and LSTM neural networks to enhance the accuracy of time series forecasting [9]. By leveraging ARIMA's proficiency in capturing linear trends and seasonality alongside LSTM's ability to model complex nonlinear patterns and long-term dependencies, the hybrid model achieves superior performance compared to either model used in isolation. This hybrid approach capitalizes on the complementary strengths of both techniques, resulting in more robust predictions that better capture the intricacies of the underlying data.

IV. PERFORMANCE METRICS

The performance evaluation of various predictive models for Bitcoin price forecasting is a critical aspect of understanding their applicability and reliability in financial markets. In this study, we conducted a comprehensive analysis of four prominent models: Auto ARIMAX, Facebook Prophet, XGBoost, and Long Short-Term Memory (LSTM) networks. Our analysis involved assessing each model's accuracy using key performance metrics such as [20] Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), alongside evaluating their overall predictive power through a novel accuracy percentage calculation. Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are commonly used metrics to evaluate the accuracy of predictive models. RMSE measures the average magnitude of the errors between predicted and actual values, giving more weight to large errors due to the square term. On the other hand, MAE represents the average absolute difference between predicted and actual values, providing a more straightforward measure of accuracy. Lower values of RMSE and MAE indicate better model performance. Which is calculated for each forecasting model such as ARIMAX, FB Prophet, XGBoost, and LSTM.Additionally, R-squared (R2) is a statistical measure that represents the proportion of variance in the dependent variable (target) that is explained by the independent variables (features) in a regression model. R2 values range from 0 to 1, where 1 indicates that the model explains all the variability of the response data around its mean.Higher R2 values signify a better fit of the model to the data, indicating its predictive power.which is specifically used for the XGBoost model.

Table 1: Performance metrics table.

Model	RMSE	MAE	R2
ARIMAX	18100.1	10733.4	-
FB Prophet	1300.2	684.8	-
LSTM	947.95	898624.02	-

XGBOOST	13417.9	6401.7	0.1628
FB Prophet + XGBOOST	886.9	523.9	-

Table 1 displayed RMSE, MAE and R2 value of the models of this study. The performance metrics table indicates that the RMSE values of the models(ARIMAX, FB Prophet, LSTM, XGBOOST and FB Prophet + XGBOOST) were 18100.1, 1300.2, 947.95, 13417.9 and 886.9 respectively and the MAE values of the models(ARIMAX, FB Prophet, LSTM, XGBOOST and FB Prophet + XGBOOST) were 10733.4, 684.8, 898624.02, 6401.7 and 523.9 respectively and the R2 value of the XGBOOST model were 0.1628.

V. RESULTS

Using FB Prophet algorithm we have achieved that forecast price against actual price effectively at certain period of time**.**

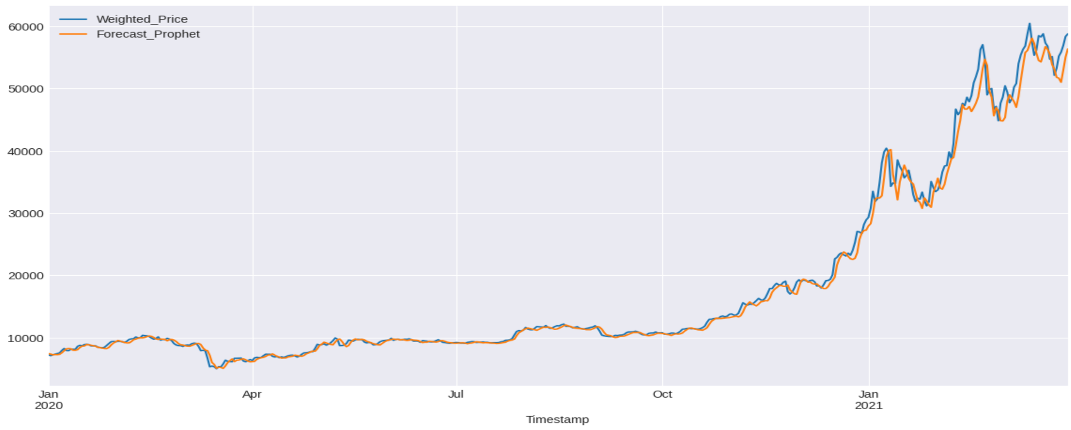

Figure 2. FB Prophet Forecast vs Actual price.

After undergone various tuning we have achieved that forecast price against actual price better by using XG_Boost at certain period of time.

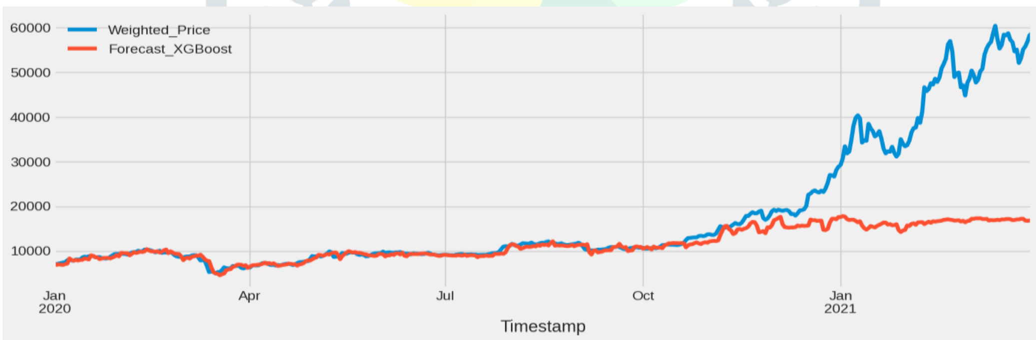

Figure 3. XGBOOST Forecast vs Actual price.

Using LSTM algorithm we have achieved that forecast price against actual price effectively at certain period of time.

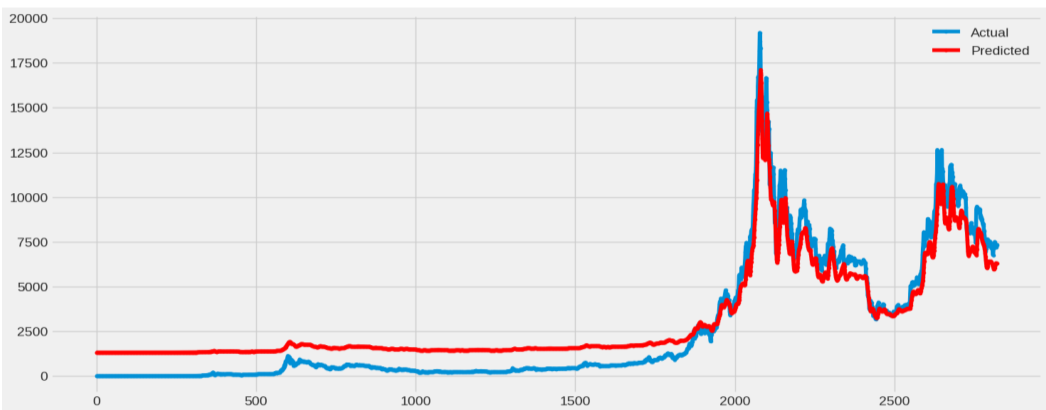

Figure 4. LSTM Forecast vs Actual price.

By comparing four models we have concluded that the FB Prophet outperforms than the three model for the given dataset.

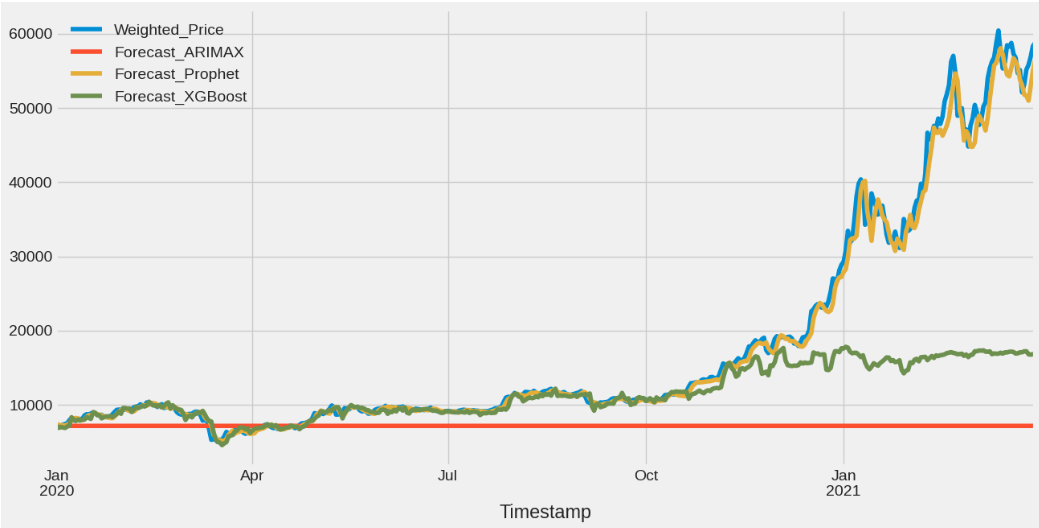

Figure 5. Weighted price vs ARIMAX vs FB Prophet vs XGBOOST.

Combining all the four model we have trained it according to the given dataset and forecast the price.By analyzing we concluded that hybriding all the four model not so effective than the other individual model and other hybrided model which experimental results given below.

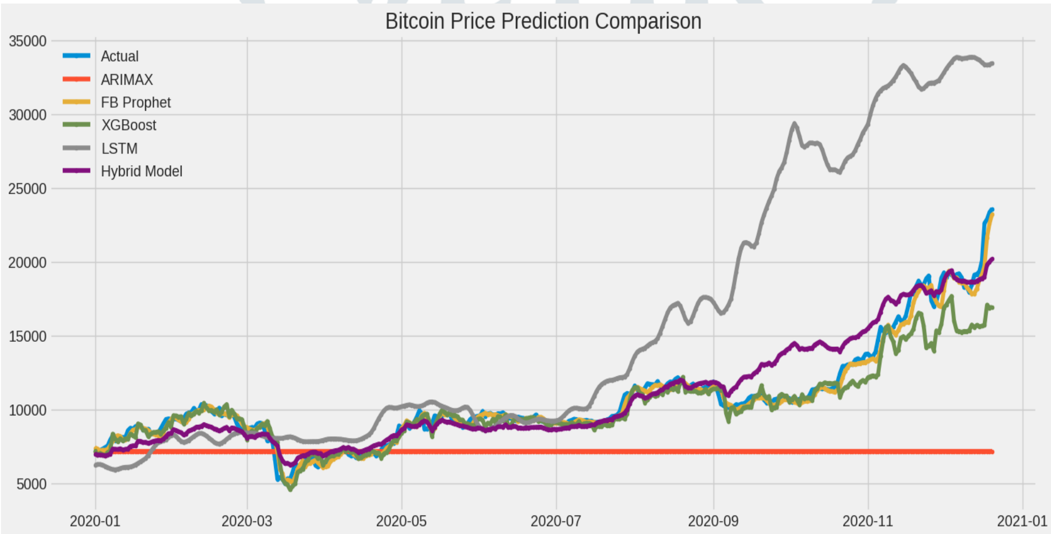

Figure 6. Hybrid model vs Actual price.

By combining the Fb Prophet and XG Boost we noticed that it performs better than the above hybrid model, but it is not so effective than the Fb Prophet model and LSTM model.

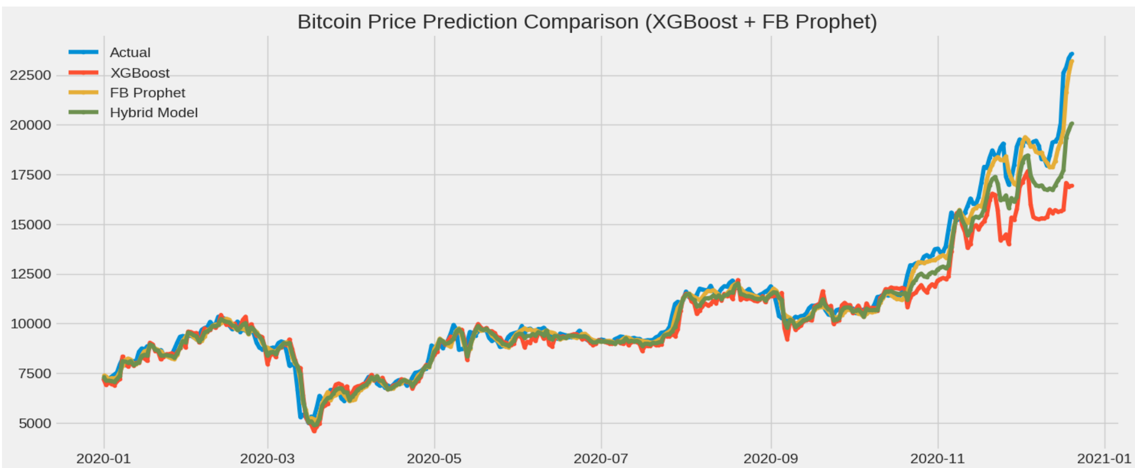

Figure 7. Hybrid model (FB prophet and XGBOOST) Forecast vs Actual price.

VI. CONCLUSION AND FUTUREWORK

In conclusion, we explored the predictive capabilities of various models to forecast Bitcoin prices. We employed ARIMAX, Facebook Prophet, XGBOOST, and LSTM models to predict the prices of Bitcoin. The implementation of a FB Prophet model, LSTM and hybrid model combining FB Prophet and XG Boost algorithms demonstrates promising results in time-series forecasting tasks than other model such as XG BOOST and ARIMAX. XGBOOST also performs effective than ARIMAX and in certain period of time after undergoing various tuning.Through the integration of both linear and nonlinear modeling techniques, the hybrid model effectively captures diverse patterns and trends present in the data, leading to improved forecasting accuracy compared to some individual models such as ARIMAX and XGBOOST. The hybrid approach leverages the strengths of each algorithm, with FB Prophet adept at modeling linear trends and XGBOOST excelling in capturing complex temporal dependencies. By combining these complementary capabilities, the hybrid model achieves enhanced predictive performance, It may be a valuable tool for various applications requiring accurate time-series forecasting after iterative tuning.Furthermore, the experimental evaluation conducted on real-world datasets validates the effectiveness of the Facebook Prophet, a forecasting tool developed by Facebook, demonstrated superior performance among the traditional models with a MAE value of 684.8. Its ability to handle various trends and seasonality in the data contributed to its effectiveness in Bitcoin price prediction. And LSTM, a type of recurrent neural network, exhibited remarkable accuracy in predicting Bitcoin prices with a RMSE value of 947.95. Its capability to capture temporal dependencies and learn intricate patterns in the data contributed to its outstanding performance.Both performed well than other model in different domains and time-series characteristics.The comparative analysis against standalone FB Prophet and LSTM models illustrates that hybrid approach performs lesser than these model with the RMSE value of 886.9 in terms of forecast accuracy and robustness. Notably, the hybrid model may exhibits superior performance in scenarios with both short term fluctuations and long-term trends after various tuning process,it would showcasing its versatility and adaptability to diverse forecasting challenges. Then these findings would underscore the practical significance of adopting hybrid modeling techniques for improving the reliability and precision of time-series predictions in various domains, including finance, energy, and healthcare. Overall, the FB prophet, LSTM and hybrid model combining FB Prophet and XGBOOST represents a promising approach in time-series forecasting, with potential for continued refinement and application in diverse domains, thereby contributing to advancements in predictive analytics and decision-making processes.

Moving forward, future research directions may focus on further enhancing the hybrid model's performance through advanced optimization techniques, feature engineering strategies, and model architecture modifications. Additionally, exploring the integration of other forecasting algorithms or ensembling methods could offer additional insights into improving predictive accuracy and robustness. Furthermore, investigating the scalability and computational efficiency of the hybrid model for large-scale time series datasets and considering various factors affecting cryptocurrency market remains an important area for exploration, especially in the context of real-time applications and big data environments. And there are plans to create an interactive, user-friendly web platform. This platform will prioritize user engagement, real-time updates, and Additional advancements comprise integrating additional data sources for a more thorough examination and offering comprehensive instructional materials within the interface to help users in understanding Bitcoin market nuances and model limitations.

REFERENCES

[1] S. Nakamoto. (2009). Bitcoin: A Peer-to-Peer Election Caash System.CryptographyMailing List.[online].Available: https://metzdowd.com.

[2] Blockchain for Dummies (2nd ed.) by Dummies Technology Experts (2019). John Wiley & Sons.

[3] Hongze Guo., Ke Gao., Yue Yu., Yingchang Liu., Lei Fu., A Diluted Bitcoin-Dollar-Gold Mean Prediction Scheme Based on Periodic Prediction Method in Institute of Electrical and Electronics Engineers,2022.

[4] Nisarg P. Patel., Raj Parekh., Nihar Thakkar., Rajesh Gupta., Sudeep Tanwar., Gulshan Sharma.,Fusion in CryptocurrencyPrice Prediction: A Decade Surveyon Recent Advancements, Architecture, and Potential Future Directionsin Institute of Electrical and Electronics Engineers,2022.

[5] Gyeongho Kim., Dong-Hyun Shin., Jae Gyeong Choi., Sunghoon Lim.A Deep Learning-Based Cryptocurrency Price Prediction Model That Uses On-Chain Datain Institute of Electrical and Electronics Engineers,2022.

[6] Firat Akba, Ihsan Tolga Medeni, Mehmet Guzel., Manipulator Detecting in Cryptocurrency Markets Based on Forecasting Anomalies,2021.

[7] Sattarov Otabek; Jaeyoung Choi.TwitterAttribute Classification With Q-Learning on Bitcoin Price Prediction in Institute of Electrical and Electronics Engineers,2022.

[8] Sarat Chandra Nayak., Subhranginee Das., Satchidananda Dehuri., Sung-Bae Cho. An Elitist Artificial Electric Field Algorithm Based Random Vector Functional Link Network for Cryptocurrency Prices Forecasting in Institute of Electrical and Electronics Engineers,2023.

[9] Xiaoxu Du., Zhenpeng Tang., Junchuan Wu., Kaijie Chen., Yi Cai.A New Hybrid Cryptocurrency Returns Forecasting Method Based on Multiscale Decomposition and an Optimized Extreme Learning Machine Using the Sparrow Search AlgorithmPrediction in Institute of Electrical and Electronics Engineers,2022.

[10] Jaehyun Park., Yeong-Seok Seo.Twitter Sentiment Analysis-Based Adjustment of Cryptocurrency Action Recommendation Model for Profit MaximizationPrediction in Institute of Electrical and Electronics Engineers,2023.

[11] Binjie Chen., Fushan Wei., Chunxiang Gu.Bitcoin Theft Detection Based on Supervised Machine Learning Algorithms in Security and Communication Networks, 2021.

[12] Firat Akba, Ihsan Tolga Medeni,Mehmet Serdar Guzel, Iman Askerzade. Manipulator Detection in Cryptocurrency Markets Based on Forecasting Anomalies in Institute of Electrical and Electronics Engineers,2021.

[13] RajParekh, Nisarg P. Patel, Nihar Thakkar, Rajesh Gupta, Sudeep Tanwar, Gulshan Sharma.DL-Guess: Deep Learning and Sentiment Analysis-Based Cryptocurrency Price Predictionin Institute of Electrical and Electronics Engineers,2021.

[14] Noor.Nayyer,NadeemJavaid, MariamAkbar, AbdulazizAldegheishem, Nabil Alrajeh, Mohsin Jamil.A New Framework for Fraud Detection in Bitcoin Transactions Through Ensemble Stacking Modin Institute of Electrical and Electronics Engineers,2023.

[15] Chang-Yi Lin., Hsiang-Kai Liao., Fu-Ching Tsai.A Systematic Review of Detecting Illicit Bitcoin Transactions in Sciencedirect,2022.

[16] Wohold, P., Brecht, M., Neumann, M., & Yashar, A. (2020). Financial Time Series Forecasting with Machine Learning. In Financial Big Data: Deep Learning and Machine Learning for Asset Management (Y. Yao, & J. Zhang, Eds.), Apress, pp. 23-54. https://doi.org/10.1007/978-1-4842-5688-1_2.

[17] Hyndman, R. J., & Athanasopoulos, G. (2013). Forecasting: principles and practice (2nd ed.). OTexts. https://otexts.com/fpp2/.

[18] Reisman, A., Bleich, A., & Dvir, I. (2018). Using XGBoost to Analyze Stock Splits. The Journal of Portfolio Management, 44(2), 112-122.

[19] Sean J. Taylor, Benjamin Letham, Arif Nabi (2018). Forecasting at Scale with Prophet. https://peerj.com/preprints/3190.pdf.

[20] Willmott, C. J. (1981). On the validation of models.Physical Geography,2(2),184194. https://www.tandfonline.com/doi/abs/10.1080/02723646.1981.10642213.

[21] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control (5th ed.). John Wiley & Sons.

[22] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics) (2nd ed.). Springer New York.

[23] Data Collection: What It Is, Methods & Tools + Examples"by QuestionPro. https://www.questionpro.com/help/survey-question-and-answer-types.html.

[24] Cryptoasset Volatility: A Primer" by International Monetary Fund(IMF) https://www.imf.org/en/Blogs/Articles/2023/07/18/crypto-needs-comprehensive-policies-to-protect-economies-and-investors.