A Comparative Review of Machine Learning Approaches for Bitcoin Price Prediction: Bridging the Gap with Deep Learning

Preeti Bisen[1], Ankita Tiwari[2], Nitya Khare[3]
[1]*M.Tech. Scholar, Sagar Institute of Research & Technology – Excellence*
[2,3]*Assistant Professor, Sagar Institute of Research & Technology – Excellence*

Abstract-The volatile nature of cryptocurrency markets, especially Bitcoin, has driven significant research interest in developing predictive models using Machine Learning (ML) and Deep Learning (DL). This review paper examines the application of ML algorithms such as Linear Regression and Random Forest in forecasting Bitcoin prices. Despite their simplicity and interpretability, these models often fall short in capturing the complex temporal dependencies and nonlinear dynamics of cryptocurrency data. We propose a hybrid Deep Learning-based approach to improve prediction accuracy and present a future direction for research in this field.

Keywords: Bitcoin, Machine Learning, Linear Regression, Random Forest, Cryptocurrency Forecasting, Deep Learning, Time Series Prediction

I.INTRODUCTION

Bitcoin, the most prominent cryptocurrency, has demonstrated extreme volatility driven by factors such as investor sentiment, regulatory developments, macroeconomic indicators, and technological advancements. Accurately forecasting Bitcoin prices has become a critical area of research in financial data science.

Traditional statistical models like ARIMA and Linear Regression often fail to perform well due to their assumptions about linearity and stationarity, which do not hold in real-world crypto markets. In contrast, machine learning techniques, including Linear Regression (LR) and Random Forest (RF), are capable of learning complex patterns from data. However, they also struggle with temporal dependencies inherent in time-series data.

To overcome these limitations, researchers have begun exploring Deep Learning models like LSTM and GRU that can capture long-term dependencies. This paper aims to review the existing literature on ML-based prediction methods and propose a DL-based model for more accurate forecasting.

II.LITERATURE REVIEW

This section presents a detailed analysis of notable studies from 2018 to 2025 that used ML techniques, particularly Linear Regression and Random Forest, for predicting Bitcoin prices.

1. Zhu et al. (2023)

This study conducted a comparative analysis of Ordinary Least Squares (OLS), Random Forest (RF), LightGBM, and Long Short-Term Memory (LSTM) networks. Using a dataset of daily Bitcoin prices, the authors found that OLS and RF performed well under low volatility conditions but failed during sudden market movements. LSTM showed robustness in learning from long-term dependencies, outperforming all other models in terms of RMSE and MAE [1].

2. Pan et al. (2023)

Pan and colleagues examined three different forecasting models — ARIMA, RF, and LSTM — across multiple cryptocurrencies, including BTC, ETH, and LTC. While RF and ARIMA had reasonable performance, they failed to capture nonlinear trends. LSTM models demonstrated superior prediction accuracy, especially for multi-step forecasting. Their work emphasized the necessity of sequence modeling in the cryptocurrency domain [2].

3. Patel & Shah (2022)

This paper utilized Random Forest and Support Vector Machine (SVM) models to predict Bitcoin price movements using technical indicators like MACD, RSI, and Bollinger Bands. RF models showed more stable and accurate results, while SVM struggled with highly volatile instances. The authors concluded that ensemble models like RF could handle complex, high-dimensional data better than simpler models like SVM [3].

4. Nguyen et al. (2021)
Nguyen and his team assessed the performance of various ML models, including LR, K-Nearest Neighbors (KNN), RF, and Gradient Boosting Machines (GBM). Their findings revealed that RF handled nonlinear relationships better than LR and KNN, although all models underperformed during periods of high volatility. This paper highlighted the need for temporal feature extraction [4].

5. Al-Ameen et al. (2021)
This research focused on combining sentiment analysis with RF models for short-term Bitcoin forecasting. Social media data was extracted and sentiment scores were incorporated as input features. RF models improved accuracy significantly with this additional feature. The study demonstrated that integrating external, real-time data can enhance predictive power [5].

6. Singh & Kaur (2020)
Singh and Kaur explored the use of LR, Ridge, and Lasso regression models in cryptocurrency prediction. Their evaluation revealed that linear models performed inconsistently due to the market's non-stationary behavior. The authors argued for the adoption of non-linear models or models capable of capturing higher-order interactions [6].

7. Hasan et al. (2019)
This paper presented a comparative evaluation of LR and RF for Bitcoin price prediction. By incorporating technical indicators like EMA, SMA, and OBV, RF demonstrated a higher accuracy rate (~82%) compared to LR (~65%). The authors concluded that RF's ability to model complex relationships made it more suitable for this domain [7].

8. Zhang et al. (2019)
Zhang applied Principal Component Analysis (PCA) for dimensionality reduction, followed by LR and RF models. PCA helped eliminate noise from the dataset, improving RF accuracy to 87%. LR, however, failed to generalize well. The study emphasized the importance of feature selection in improving ML model performance [8].

9. Yadav & Mishra (2018)
This study used LR with multiple technical indicators and moving averages to forecast Bitcoin prices. Despite tuning hyperparameters, the model struggled to predict sudden price spikes. The authors concluded that the linear model's rigidity made it unsuitable for highly volatile datasets [9].

10. Rahman et al. (2020)
Rahman proposed an ensemble learning method combining LR and RF to forecast Bitcoin prices. Although the ensemble improved robustness and reduced overfitting, it increased computational complexity. The paper recommended using this method only when computational resources are not a constraint [1].

Research Gap
While ML models like LR and RF have shown potential, they lack the ability to model complex temporal dynamics, which are inherent in Bitcoin's price behavior. Many studies fail to integrate temporal-aware models such as RNNs or LSTMs, and few benchmark ML against DL on a unified dataset. Additionally, most works focus on short-term predictions and do not explore multi-step forecasting.

III.TECHNOLOGIES USED IN LITERATURE

- Linear Regression: Simple, interpretable, but fails to handle non-linear and volatile data.
- Random Forest: Captures non-linear patterns well but lacks temporal awareness.
- Technical Indicators: Enhance feature space but do not inherently improve model structure.
- Sentiment Analysis: Improves short-term prediction by integrating social media signals.

IV. PROPOSED METHODOLOGY

We propose a hybrid Deep Learning-based forecasting model combining LSTM and GRU architectures.

Dataset
We will use publicly available OHLC (Open, High, Low, Close) Bitcoin datasets from Yahoo Finance or Kaggle, enriched with indicators like RSI, EMA, and MACD.

Preprocessing

- Handling missing values and outliers
- Normalization (Min-Max scaling or StandardScaler)
- Feature engineering: lag values, rolling statistics
- Time series windowing

Models to Compare

- Baseline: Linear Regression and Random Forest
- Proposed: LSTM and GRU (standalone and hybrid models)

Evaluation Metrics

- Root Mean Square Error (RMSE)
- Mean Absolute Error (MAE)
- R^2 Score

Tools

- Python (Jupyter, Google Colab)
- Libraries: TensorFlow, Keras, Scikit-learn, Pandas, Matplotlib

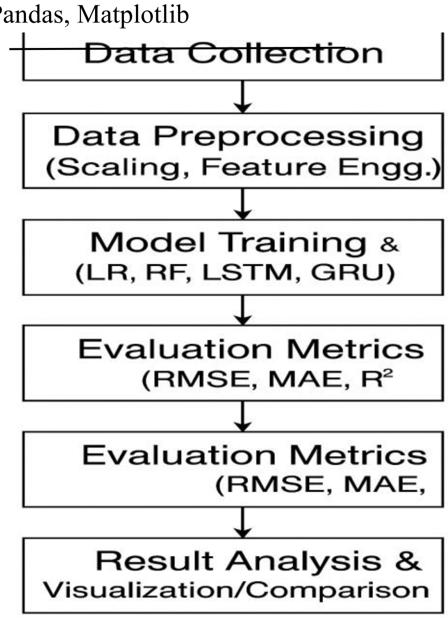

Figure 1. Hypothetical workflow of Proposed methodology

V. CONCLUSION

ML models like LR and RF provide a foundation for Bitcoin price forecasting but fall short in modeling temporal dependencies and complex non-linear patterns. DL models, especially LSTM and GRU, are better suited for such tasks. This review suggests that a hybrid DL approach could significantly enhance predictive accuracy and recommends future work in benchmark comparisons using unified datasets.

REFERENCE

[1] Zhu, G. (2023). Bitcoin Return Prediction using OLS, RF, LightGBM, and LSTM.

[2] Pan, R. (2023). Cryptocurrency Price Prediction Using ARIMA, RF, and LSTM.

[3] Patel, S., & Shah, R. (2022). Bitcoin Price Forecasting Using Random Forest and SVM.

[4] Nguyen, H. et al. (2021). Machine Learning in Cryptocurrency Forecasting.

[5] Al-Ameen, Z. et al. (2021). Sentiment-Driven Bitcoin Forecasting Using RF.

[6] Singh, R., & Kaur, G. (2020). Regression Techniques in Crypto Prediction.

[7] Hasan, T. et al. (2019). Comparative Study of RF and LR in BTC Forecasting.

[8] Zhang, L. et al. (2019). ML-Based Cryptocurrency Price Prediction Framework.

[9] Yadav, V., & Mishra, R. (2018). Short-Term Forecasting with Regression.

[10] Rahman, F. et al. (2020). Forecasting Bitcoin Prices Using Ensemble ML.