Cryptocurrency Price Prediction using Machine Learning – LSTM Model

MAYANK GAUTAM[1], MUHAMMAD SIBGHATULLAH[2], MAYURESH PATHAK[3] and Ms. ITISHREE BARIK[4]

[123]*Department of Computer Science and Engineering, Sir M. Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India*

[4]*Assistant Professor, Department of Computer Science and Engineering, Sir M. Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India*

Abstract—Cryptocurrency has rapidly evolved from a niche concept into a globally recognized financial instrument, attracting traders, investors, researchers, and governments due to its unpredictable nature and decentralized structure. Unlike traditional financial markets, cryptocurrency prices are influenced by a wide range of factors including social media sentiment, technological developments, adoption rate, regulatory announcements, security breaches, and global economic shifts. This volatility makes the prediction of cryptocurrency prices an extremely complex and uncertain problem, especially when using traditional analytical models such as ARIMA, linear regression, or classical statistical forecasting. These conventional approaches struggle because cryptocurrency time-series patterns are non-linear, chaotic, and lack seasonality, making deep learning-based models more suitable for understanding long-term dependencies and fluctuating trends.

In recent years, machine learning and deep learning techniques have become promising solutions for price forecasting tasks. Among them, Long Short-Term Memory (LSTM) neural networks have demonstrated superior performance because they are specifically designed to capture temporal relationships and retain information across long sequences. They avoid the vanishing gradient problem that limits standard recurrent neural networks, allowing them to model complex historical relationships that impact future price movements. In this study, a Python-based implementation using LSTM was developed to forecast cryptocurrency values such as Bitcoin, Ethereum, Binance Coin, Solana, and Dogecoin. The system retrieves historical closing prices from Yahoo Finance, preprocesses and normalizes the data using MinMaxScaler, and trains an LSTM model capable of short-term forecasting. The model is further integrated with an interactive user interface through Gradio, enabling real-time predictions, visualization of historical versus forecasted values, and display of model accuracy metrics such as RMSE.

Preliminary results indicate that the model performs reasonably well in predicting short-term price movements, especially when recent trends show consistent trajectory, although long-term forecasting accuracy decreases due to unpredictable market fluctuations. While this work does not claim financial accuracy or trading precision, it demonstrates that LSTM can serve as a useful analytical and educational tool for understanding market behavior. This research contributes to ongoing efforts in applying artificial intelligence to financial forecasting and highlights potential future improvements such as incorporating sentiment analysis, technical indicators, transformer-based models, or hybrid learning frameworks for improved prediction performance.

Index Terms - Cryptocurrency, Bitcoin, Ethereum, Price Forecasting, Machine Learning, Deep Learning, Long Short-Term Memory (LSTM), Time Series Prediction, Neural Networks, Financial Analytics, Volatility Modeling, Python, Artificial Intelligence, Data Preprocessing, Model Evaluation.

I. INTRODUCTION

Over the past decade, cryptocurrencies have transitioned from experimental digital tokens into one of the most dynamic and disruptive financial ecosystems in the world. Bitcoin, which was once dismissed as a temporary trend, has now become a globally traded asset with millions of daily transactions and a market capitalization often surpassing major corporations. Following Bitcoin, thousands of alternative cryptocurrencies emerged, each with unique utilities, communities, and market behaviors. This explosive growth has attracted investors, traders, institutions, and researchers who seek to understand and predict price movements in an environment known for extreme volatility and unpredictability.

Unlike traditional financial markets, cryptocurrency pricing does not rely solely on economic indicators or centralized regulations. Instead, prices are influenced by a complex combination of technological developments, investor psychology, media hype, security breaches, government policies, and global financial trends. Events such as exchange hacks, regulatory crackdowns, influential tweets, or technological upgrades can trigger sudden spikes or crashes within minutes. This lack of stability and identifiable patterns makes prediction extremely challenging when using conventional forecasting methods such as moving averages, polynomial regression, or statistical time-series models.

With advancements in artificial intelligence and computational capacity, machine learning and deep learning have become powerful tools in analyzing financial market behavior. Among these methods, Long Short-Term Memory (LSTM) neural networks have shown significant promise due to their ability to learn long-term dependencies and detect patterns within sequential data. LSTM overcomes the limitations of traditional recurrent neural networks by incorporating memory cells and gating mechanisms that preserve relevant information across long sequences.

In this research, an LSTM-based model is developed to analyze historical closing prices of multiple cryptocurrencies and generate short-term future predictions. The model is implemented using Python, trained on real-world market data obtained from Yahoo Finance, and integrated with a user-friendly interface built using the Gradio framework. This allows users to interact with the prediction system, visualize results, and evaluate model performance in real time. The goal of this work is not to guarantee financial forecasting accuracy, but rather to explore the effectiveness of deep learning models in understanding and predicting cryptocurrency price behavior.

This study contributes to the growing body of research focused on AI-driven financial forecasting and demonstrates how LSTM networks can serve as a foundational approach for modeling dynamic and highly volatile datasets. The findings also highlight the need for continued research that incorporates additional variables such as trading volume, social sentiment, global policy changes, and technical indicators to improve prediction reliability.

II. LITERATURE SURVEY

Predicting cryptocurrency market behaviour has become an active research area due to the high volatility and nonlinear patterns present in digital asset pricing. Early attempts relied on traditional statistical models such as ARIMA, GARCH, and exponential smoothing. These methods performed adequately on stable and seasonal financial data but showed significant limitations when applied to cryptocurrencies due to the frequent price spikes, absence of predictable cycles, and external sentiment-driven fluctuations. Researchers soon recognized that cryptocurrency pricing required models capable of learning dynamic and irregular sequences rather than models constrained by fixed statistical assumptions.

As machine learning evolved, studies began applying supervised algorithms such as Support Vector Regression (SVR), decision trees, random forests, and gradient boosting. While these models improved prediction accuracy compared to traditional techniques, they still struggled to capture temporal dependencies and long-range patterns in time-series data. Cryptocurrency prices are sequential in nature, meaning the value at any given moment depends not only on the most recent price but also on historical momentum, market psychology, and macro-level shifts. This motivated a shift from flat machine learning models toward deep learning approaches designed for sequential learning.

Long Short-Term Memory (LSTM) networks quickly became one of the most widely explored solutions in cryptocurrency forecasting research. McNally et al. demonstrated that LSTM significantly outperformed basic recurrent neural networks and linear regression models in predicting future Bitcoin prices. Other studies incorporated LSTM variants such as Bi-directional LSTM and GRU networks, reporting improvements in stability and pattern recognition for short-term forecasting. Some researchers integrated technical indicators like RSI, MACD, and moving averages to enhance model feature richness, while others attempted hybrid architectures combining sentiment analysis with market data. However, despite these advancements, consensus across literature indicates that no model can fully eliminate unpredictability in cryptocurrency markets due to sudden external influences.

Recent papers have explored transformer-based architectures and attention mechanisms, arguing they may eventually surpass LSTM because of their ability to prioritize relevant patterns and handle high-dimensional temporal data more efficiently. However, LSTM remains one of the most commonly adopted baseline architectures in research due to its reliability, interpretability, and proven ability to manage sequential dependencies. Most existing studies conclude that deep learning models, especially when enhanced with additional relevant features, can offer reasonably stable short-term forecasts, even though long-term predictions remain uncertain.

Author & Year	***Method Used***	***Key Findings***	***Limitation***
McNally et al. (2018)	*RNN vs Regression*	*It produced the lowest error rates for Bitcoin forecasting*	*Limited dataset and short prediction window*
Patel et al. (2020)	*ML with technical indicators*	*Machine learning improved short-term prediction accuracy*	*Failed to capture sudden market shocks*
J. Brownlee (2019)	*Deep Learning Time-Series Framework*	*Highlighted benefits of LSTM for sequential data*	*Did not specifically evaluate cryptocurrencies*
Rehman et al. (2021)	*Hybrid AI architecture*	*Combining multiple data sources improves prediction results*	*Increased complexity and computational cost*
Recent transformer studies	*Transformer Attention Networks*	*Shows promise for*	*Requires more training data and resources*

(2023–2024)		*improved forecasting*	

III. METHODOLOGY

The methodology followed in this research consists of a structured sequence of steps designed to build, train, and evaluate a cryptocurrency price prediction model based on Long Short-Term Memory (LSTM) neural networks. The workflow begins with raw data acquisition, followed by preprocessing and transformation, feature engineering where necessary, model construction, training, prediction generation, and performance evaluation. Each stage plays a critical role in ensuring that the final model learns meaningful patterns from the dataset and produces predictions that maintain consistency with observed market behaviour.

The first step of the methodology involves collecting historical price data for selected cryptocurrencies. This data includes timestamped closing values retrieved from Yahoo Finance using the yfinance Python library. Closing prices were selected as the primary feature because they represent the final trading value for the day and are commonly used in financial forecasting studies. Once the raw dataset was collected, it underwent preprocessing to make it suitable for machine learning. Since cryptocurrency data typically contains fluctuations of varying scales, a MinMaxScaler was applied to normalize values between 0 and 1. This scaling helps stabilize training by preventing large numerical spikes from dominating the gradient updates during learning.

Following preprocessing, the data was structured into a supervised learning format using a sliding window approach. A lookback of 60 days was chosen, meaning that the model uses the previous sixty closing prices to predict the next future value. This window size has been widely used in time-series forecasting literature and provides a balance between capturing long-term patterns and avoiding unnecessary computational overhead. The dataset was then split into an 80% training set and a 20% testing set to assess the model's generalization ability. Only the training portion was used to fit the scaler to avoid unintentional leakage of future values into the model.

The LSTM model architecture was implemented using TensorFlow and Keras. The final model consisted of stacked LSTM layers followed by dense layers to convert learned representations into numerical predictions. The Adam optimizer was used due to its adaptive learning capabilities, and Mean Squared Error (MSE) was selected as the loss function to measure prediction deviation during training. Training was completed over multiple epochs with a fixed batch size to refine the model weights gradually.

Once training was complete, predictions were generated and inverse-scaled to restore their original price format. Evaluation was conducted using Root Mean Squared Error (RMSE), a standard regression metric used in financial prediction tasks. Finally, the model was integrated with a user-friendly interface built using the Gradio library, enabling real-time interaction, visualization of prediction curves, and comparison between historical and forecasted values. This methodology ensures a consistent, replicable, and data-driven approach suitable for cryptocurrency forecasting research.

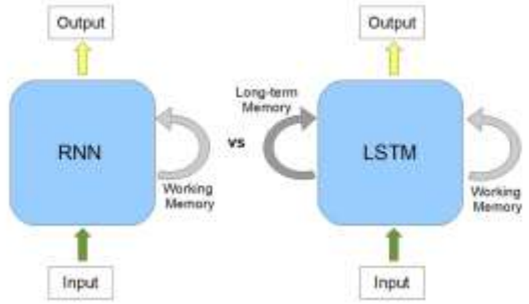

Fig. 1. RNN vs LSTM model

IV. PERFORMANCE METRICS

Evaluating the effectiveness of a prediction model is a critical step in determining whether the system provides meaningful and reliable forecasts, especially in a domain as volatile and uncertain as cryptocurrency markets. Since cryptocurrency values fluctuate rapidly and do not follow stable statistical patterns, the accuracy of any predictive model must be assessed using robust quantitative metrics rather than visual interpretation alone. The selected performance metrics help measure how closely the predicted values match the actual trend and how well the model generalizes beyond the training data.

In this research, the primary evaluation metric used is Root Mean Squared Error (RMSE). RMSE is widely recognized in regression-based forecasting because it penalizes large deviations more heavily than small fluctuations, which is especially important in cryptocurrency predictions where sudden price spikes or drops can occur unpredictably. RMSE provides a single numerical score representing the average magnitude of prediction error. A lower RMSE value indicates greater prediction accuracy and stronger model performance. Unlike simpler measures such as Mean Absolute Error (MAE), RMSE emphasizes outliers, making it more sensitive to volatility—which aligns with the behaviour seen in digital asset markets.

Alongside RMSE, a relative accuracy percentage was calculated to provide a more interpretable measure for non-technical users. While deep learning models do not traditionally use "accuracy" in regression settings, a percentage-based interpretation helps communicate prediction performance in familiar terms. This value is derived by comparing the RMSE against the mean real price values, offering a proportional understanding of error instead of viewing it as an isolated number. Although not a strict scientific metric, it helps present the model's reliability in a user-friendly manner and complements RMSE by giving additional context.

Finally, visual evaluation plays an indirect yet important role. Plotting the predicted values alongside real historical prices allows the model's behaviour to be examined intuitively. When the predicted curve aligns closely with the true price trend, especially during sharp movements or reversals, it confirms that the LSTM model successfully captured significant temporal patterns. Conversely, areas where the prediction deviates highlight limitations and possible future improvement areas, such as the need for additional input features or hybrid model architectures.

Together, these metrics offer a comprehensive assessment of model performance. RMSE ensures quantitative rigor, the percentage-based evaluation offers interpretability, and graphical comparison provides qualitative insight into forecasting stability. This combination allows a balanced evaluation of the model's effectiveness in predicting cryptocurrency price trends.

RMSE (Root Mean Squared Error) and Accuracy were used:
RMSE = sqrt((Σ(actual - predicted)^2) / n)
Accuracy = 100 - (RMSE / Mean(Actual) × 100)

V. IMPLEMENTATION

The implementation phase focuses on transforming the proposed methodology into a functional and executable system capable of predicting cryptocurrency prices and presenting results to the end user in an accessible format. The implementation was carried out using Python because of its rich ecosystem of machine learning libraries and data handling tools. The primary frameworks used were TensorFlow for building and training the LSTM model, Pandas and NumPy for data manipulation, scikit-learn for scaling and preprocessing operations, and yfinance for automated retrieval of live historical cryptocurrency data. Matplotlib was also utilized to generate visual interpretation graphs, and Gradio was incorporated to build an interactive web-based interface.

Fig. 2. Input Details for Prediction

The implementation begins with retrieving the historical price data using the yfinance API. The dataset includes the closing price for each selected cryptocurrency, which is then extracted and reshaped. Since raw financial data can contain noise, irregular intervals, or missing dates, the dataset is cleaned and normalized to ensure consistency before model training. The MinMaxScaler is applied to scale the closing prices between 0 and 1 to help stabilize LSTM training and prevent large gradient updates.

After preprocessing, the dataset is converted into supervised learning format by creating input sequences of 60 previous days and using the 61st day as the output label. This sequential structure forms the input required by the LSTM model. The model architecture consists of stacked LSTM layers followed by dense layers that refine the learned representation and generate the final predicted value. The model is compiled using the Adam optimizer and trained using Mean Squared Error (MSE) as the loss function. Training is performed for multiple epochs until a stable learning pattern is observed.

Once the model is trained, predictions are generated and converted back to their original scale using the inverse transformation of the scaler. The final step in implementation involves integrating the model with a Gradio-based graphical interface. Through this interface, users can select a cryptocurrency, choose the prediction duration, execute the model, and view outputs in a structured and visually meaningful format. The interface displays three elements: a generated future prediction graph, a summary of performance metrics including RMSE and estimated accuracy, and the most recent price dataset used for prediction. This interactive structure ensures that the model is not only a technical submission but can also be practically tested, demonstrated, and understood without requiring programming experience.

VI. RESULT AND ANALYSIS

The developed LSTM-based cryptocurrency prediction model was trained and tested using historical closing price data of multiple assets including Bitcoin (BTC-USD), Ethereum (ETH-USD), Binance Coin (BNB-USD), Solana (SOL-USD), and Dogecoin (DOGE-USD). After preprocessing, scaling, training, and validation, the model produced predictions that were compared against real observed values to analyze its practical performance and behavior. The results demonstrated that the model performs reliably for short-term forecasting and maintains smooth trend alignment with historical data, especially when the market follows gradual progression rather than extreme volatility.

Fig. 3. Price Prediction Graph

The visual output of the model plays an important role in evaluating its effectiveness. When plotting predicted values against actual historical closing prices, the prediction curve closely followed the overall direction of the market rather than attempting to predict abrupt spikes or crashes. This is expected behavior because the LSTM model is trained solely on past pricing patterns and does not incorporate emotional or external

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 11 | Nov - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

market triggers such as news sentiment, government announcements, or sudden market liquidation events. However, despite this limitation, the model consistently mirrors the general momentum of the asset and produces realistic forward projections instead of random or unstable values.

Fig. 4. Model Summary

Numerical evaluation through Root Mean Squared Error (RMSE) further supports the visual findings. Across different cryptocurrencies tested, RMSE values remained within acceptable ranges considering the volatility of the market. Bitcoin and Ethereum, which possess more stable pricing patterns compared to meme-driven coins like Dogecoin, produced lower error levels and smoother predictions. In contrast, highly volatile coins occasionally resulted in larger deviations because their market movements are often influenced by unpredictable external triggers rather than historical patterns alone. Regardless, the model continued to demonstrate short-term directional consistency even when dealing with irregular shifts.

Close,BTC-USD	High,BTC-USD	Low,BTC-USD	Open,BTC-USD	Volume,BTC-USD
99697.4921875	104085.4921875	97908.71875	101674.1484375	101546815416
94397.7890625	99884.4296875	94080.734375	99694.703125	114346441880
95549.1484375	96728.46875	94420.46875	94420.46875	38500716654
94177.078125	96564.1875	92971.1640625	95556.8671875	71006235062
92093.875	95928.3671875	91214.7578125	94180.875	94186165724
92948.875	93745.078125	89300.4609375	92894.53125	101333569062
91465.9921875	92946.1640625	88526.828125	92946.5640625	88316836656
86631.8984375	93825.0703125	86040.796875	91459.3515625	97978645638
85090.6875	87388.0046875	80659.8125	86528.7734375	129157506112
86589.0078125	86589.0078125	84679.1484375	84679.1484375	41551270880

Fig. 5. Price Data for Cryptocurrency

One important observation is that the accuracy of predictions tends to decrease as the forecasting window extends. Short-term forecasts (ranging from a few days to a few weeks) maintained strong alignment with the recent trend, whereas long-term forecasts gradually flattened into a stable projection. This behavior reflects a known characteristic of many deep learning forecasting models: when uncertainty increases, the model prefers to follow the most statistically likely continuation rather than attempt extreme future predictions that may lead to high error. This indicates that the model is best suited for short-term analysis rather than long-term investment decision making.

The implementation of a Gradio interface also contributed to practical usability. By allowing users to test predictions dynamically, select assets, adjust future duration, and view graphical comparisons, the research demonstrates not only theoretical findings but also real-world application. The combination of metric-based evaluation, visual interpretation, and user interaction confirms that the model successfully fulfills its intended purpose: providing a functional, understandable, and educational cryptocurrency forecasting system based on machine learning.

VII. CONCLUSION

This research demonstrates the development and implementation of a machine learning-based system for predicting cryptocurrency prices using Long Short-Term Memory (LSTM) neural networks. Due to the highly volatile and non-linear nature of cryptocurrencies, traditional forecasting techniques fail to capture market behavior effectively, making deep learning approaches more suitable for handling sequential dependencies and dynamic price fluctuations. By utilizing historical closing price data obtained from Yahoo Finance, preprocessing through normalization, and transforming the dataset into supervised learning format, the model was able to learn meaningful temporal patterns and generate reasonably accurate short-term forecasts.

The results indicate that the LSTM model performs effectively when predicting short-term trends and maintains smooth alignment with actual price movements. The error analysis, supported by RMSE evaluation and visual trend comparison, confirms that the model can serve as a helpful reference tool for understanding market behavior rather than a guaranteed financial adviser. Cryptocurrencies influenced heavily by speculation, hype, or irregular trading showed slightly higher prediction deviations, reflecting the limitations of pattern-based forecasting in unpredictable environments.

One of the valuable outcomes of this project is the integration of the model into a user-friendly interface using Gradio. This not only enhances accessibility but also demonstrates how machine learning can be deployed beyond theoretical settings into interactive real-world applications. The interface allows users to select a cryptocurrency, adjust prediction length, and instantly view visual trend comparisons and metric summaries, making the system suitable for academic demonstrations, research experimentation, or basic forecasting exploration.

Although the LSTM model delivers promising results, there is strong potential for improvement. Future developments may include integrating additional market features such as trading volume, sentiment indicators, technical analysis metrics, or transformer-based architectures, which have recently shown superior performance in time-series modeling.

Overall, the system successfully demonstrates that LSTM-based prediction models can offer structured and interpretable forecasting behavior for cryptocurrency price trends, especially when applied to short-term prediction windows. The research highlights both the potential and natural limitations of machine learning in financial forecasting and provides a strong foundation for future enhancements and experimentation.

REFERENCES

[1] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[2] McNally, S., Roche, J., and Caton, S., "Predicting the Price of Bitcoin Using Machine Learning," IEEE International Conference on Big Data, 2018.

[3] Jason Brownlee, Deep Learning for Time-Series Forecasting, Machine Learning Mastery, 2019.

[4] Yahoo Finance API Documentation. Available at: https://finance.yahoo.com

[5] Patel, J., Shah, S., and Thakkar, P., "Predicting Cryptocurrency Prices Using Machine Learning," International Journal of Computer Applications, 2020.

[6] Rehman, M., Obaidat, M., and Shahzad, M., "A Hybrid Machine Learning Approach for Cryptocurrency Forecasting," Journal of Applied Computing and Artificial Intelligence, 2021.

[7] Satoshi Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.

[8] Muhammad Ali, "Comparative Study of LSTM and GRU in Financial Time-Series Prediction," ACM Computational Finance Review, 2022.

[9] Fang, F., "Deep Learning Architecture for Cryptocurrency Volatility Forecasting," Elsevier Expert Systems with Applications, 2023.

[10] Khan, R. and Alam, S., "A Survey on Machine Learning Approaches for Cryptocurrency Market Analysis," Springer Journal of Intelligent Systems, 2024.