International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 06 June-2026 | Impact Factor: 3.5

A Streaming Data Collection and Analysis for Bitcoin Using LSTM Algorithm

Sankari.S[1], S.Bhuvaneshwari[2], A. B. Hajira Be[3]

[1]PG Student, Department of Computer Applications, Karpaga Vinayaga College of Engineering and Technology, Chinna Kolambakkam, Maduranthagam Taluk, Chengalpattu District, Tamil Nadu – 603308, Gmail: sankariselven1626@gmail.com

[2]Assistant Professor, Department of Computer Applications, Karpaga Vinayaga College of Engineering and Technology, Chinna Kolambakkam, Maduranthagam Taluk, Chengalpattu District, Tamil Nadu – 603308, Gmail: bhuvaneshwari.s.mca@gmail.com

[3]Associate Professor, Department of Computer Applications, Karpaga Vinayaga College of Engineering and Technology, Chinna Kolambakkam,Maduranthagam Taluk, Chengalpattu District , Tamil Nadu – 603308, Gmail: hajiraab786@gmail.com

How to Cite this Article:

Sankari.S, & Be, A. B. H. (2026). A Streaming Data Collection and Analysis for Bitcoin Using LSTM Algorithm. International Journal of Creative and Open Research in Engineering and Management, 02(6). https://doi.org/10.55041/ijcope.v2i6.174

https://doi.org/10.55041/ijcope.v2i6.174

Abstract—

Cryptocurrency markets have gained significant global attention due to their decentralized nature and high financial value. Among various cryptocurrencies, Bitcoin is the most widely traded and exhibits highly volatile price behavior. Accurate analysis and prediction of Bitcoin price trends are challenging because the market is influenced by rapid trading activities, large data streams, and complex temporal patterns. This paper presents a streaming data collection and analysis system for Bitcoin using the Long Short-Term Memory (LSTM) deep learning algorithm. The proposed system continuously collects real-time Bitcoin market data from online cryptocurrency exchanges through streaming APIs. The collected data is then preprocessed and analyzed using an LSTM-based predictive model capable of learning long-term dependencies in time-series data. The LSTM network processes sequential historical price data to forecast future market trends and provide analytical insights into Bitcoin price movements. The system integrates data acquisition, preprocessing, deep learning-based prediction, and visualization modules to create an efficient cryptocurrency analysis framework. The proposed approach focuses on improving prediction accuracy by combining real-time streaming data with advanced neural network models. This system can assist researchers, financial analysts, and investors in understanding cryptocurrency market behavior and making informed trading decisions. The proposed design demonstrates the feasibility of integrating streaming data technologies with deep learning models for real-time financial market analysis.

Keywords— Cryptocurrency, Bitcoin, Streaming Data, LSTM Algorithm, Deep Learning, Time-Series Prediction, Financial Data Analysis.

I. INTRODUCTION

Cryptocurrency has emerged as a revolutionary digital financial technology that enables secure, decentralized, and peer-to-peer financial transactions without the need for traditional banking institutions. Among the various cryptocurrencies available today, Bitcoin remains the most popular and widely traded digital asset in global financial markets. Since its introduction in 2009, Bitcoin has attracted significant attention from investors, financial analysts, and researchers due to its rapid price fluctuations and high investment potential. However, predicting Bitcoin price movements is extremely challenging because the cryptocurrency market is highly volatile and influenced by multiple dynamic factors such as trading volume, investor sentiment, global economic events, and market speculation.

Traditional financial forecasting methods rely mainly on statistical models that assume linear relationships within the data. However, cryptocurrency price movements often exhibit complex nonlinear patterns and long-term dependencies that cannot be effectively captured using conventional approaches. As a result, machine learning and deep learning techniques have gained increasing importance in financial data analysis and prediction. Among these techniques, the Long Short-Term Memory (LSTM) neural network has proven to be highly effective for time-series prediction problems due to its ability to learn sequential patterns and retain long-term dependencies within large datasets.

Another major challenge in cryptocurrency analysis is the continuous generation of large volumes of market data from multiple exchanges and trading platforms. Real-time streaming data technologies allow systems to collect and process live financial data continuously, enabling more accurate and timely analysis. By integrating streaming data collection with advanced deep learning algorithms, it becomes possible to build intelligent systems capable of monitoring market trends and predicting future price movements more effectively.

This paper presents the design and implementation of a streaming data collection and analysis system for Bitcoin price prediction using the LSTM algorithm. The proposed system collects real-time Bitcoin market data through online APIs, performs preprocessing and feature extraction, and applies an LSTM-based deep learning model to analyze historical and streaming data sequences. The system generates predicted price trends and visual representations that help in understanding market behavior.

II.PROBLEM DEFINITION AND SOLUTION

Existing System

Currently, most cryptocurrency analysis and prediction systems rely on traditional statistical models or basic machine learning techniques. These systems generally analyze historical market data to identify patterns and forecast future price trends. Methods such as linear regression, moving averages, and autoregressive models have been widely used for financial forecasting. However, these techniques often fail to capture the complex nonlinear patterns and rapid fluctuations present in cryptocurrency markets.

In many research studies, machine learning models such as Support Vector Machines (SVM), Decision Trees, and Artificial Neural Networks (ANN) have been applied to Bitcoin price prediction. While these approaches provide improved performance compared to statistical models, they still face limitations when handling large-scale sequential financial data and long-term dependencies within time-series datasets.

Another major limitation of existing systems is the lack of efficient real-time data processing mechanisms. Many systems rely on static datasets collected over specific periods, which limits their ability to analyze continuously changing cryptocurrency markets. Without real-time data streaming, prediction models may not reflect the most recent market conditions.

Additionally, traditional neural networks struggle with remembering long-term dependencies in time-series data due to the vanishing gradient problem. This makes it difficult to accurately predict future price trends based on past sequential patterns. As a result, there is a need for more advanced deep learning techniques and streaming data processing frameworks to improve cryptocurrency analysis and forecasting performance.

Problem Definition

• Cryptocurrency markets generate massive volumes of real-time trading data that require efficient collection and processing methods.

• Traditional statistical models fail to capture complex nonlinear patterns present in Bitcoin price movements.

• Many prediction systems rely only on historical datasets without utilizing real-time streaming data.

• Conventional machine learning algorithms struggle with long-term dependencies in time-series data.

• There is a need for an intelligent system that can analyze streaming Bitcoin data and predict future market trends accurately.

Proposed Method

The proposed method introduces a streaming data collection and analysis system for Bitcoin price prediction using the Long Short-Term Memory (LSTM) deep learning algorithm. The system is designed to continuously collect real-time Bitcoin market data from cryptocurrency exchange platforms through application programming interfaces (APIs). This streaming data includes parameters such as Bitcoin price, trading volume, timestamps, and market trends.

The collected data is first processed through a preprocessing stage where noise removal, normalization, and feature extraction are performed. This step ensures that the dataset is properly formatted and optimized for deep learning analysis. The processed data is then organized into sequential time-series inputs suitable for training and testing the predictive model.

The LSTM neural network is used as the core prediction model due to its ability to learn long-term dependencies in sequential data. Unlike traditional neural networks, LSTM networks contain memory cells and gating mechanisms that allow them to retain important information over long periods. This capability makes LSTM particularly effective for analyzing financial time-series data such as cryptocurrency price movements.

The trained LSTM model analyzes historical patterns along with newly streamed data to forecast future Bitcoin prices. The system continuously updates its predictions as new data arrives, allowing real-time market analysis. Additionally, visualization tools are integrated to display predicted price trends and market insights for users.

By combining real-time data streaming with deep learning techniques, the proposed system provides a more accurate and efficient framework for cryptocurrency market analysis. The system can assist financial analysts, investors, and researchers in understanding Bitcoin market behavior and making informed trading decisions.

III. BLOCK DIAGRAM AND ITS DESCRIPTION

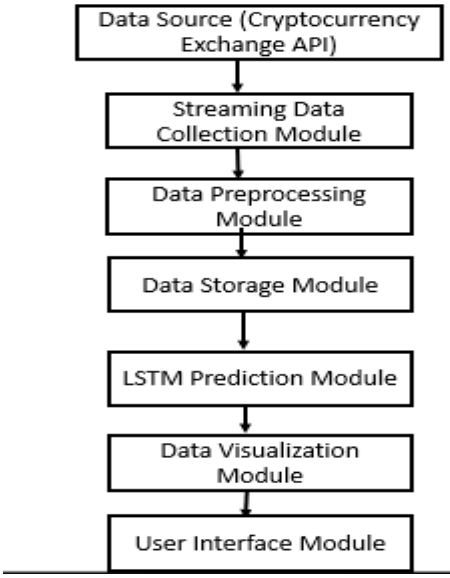

Figure 1: Proposed System Architecture for Cryptocurrency Price Prediction Using LSTM

The block diagram illustrates the overall architecture of the proposed streaming data collection and analysis system for Bitcoin price prediction using the LSTM algorithm. The system integrates multiple modules responsible for data acquisition, preprocessing, prediction, and result visualization. The entire process operates continuously to analyze real-time cryptocurrency market data and generate predictive insights.

The Data Source module represents the cryptocurrency exchange platforms that provide real-time Bitcoin market data through application programming interfaces (APIs). These APIs supply continuous streams of information such as Bitcoin

price, trading volume, timestamp, and market activity.

The Streaming Data Collection Module is responsible for capturing live Bitcoin data from the exchange APIs. This module continuously receives incoming data streams and forwards them to the preprocessing stage. The streaming mechanism ensures that the system always processes the most recent market information.

The Data Preprocessing Module performs data cleaning and transformation. It removes noise, handles missing values, and normalizes the data to make it suitable for machine learning analysis. Feature extraction is also performed to organize the dataset into structured time-series sequences required for prediction.

The Data Storage Module temporarily stores both historical and streaming data. This storage component helps maintain past market records that are required for training the deep learning model and for performing future analysis.

The LSTM Prediction Module is the core analytical component of the system. The Long Short-Term Memory neural network processes the sequential data and learns long-term dependencies within Bitcoin price movements. Based on the trained model, the system predicts future Bitcoin price trends using both historical and newly streamed data.

The Data Visualization Module converts the prediction results into graphical representations such as line charts and trend graphs. These visualizations help users understand market behavior and compare predicted prices with actual market values.

Finally, the User Interface Module provides interaction between the system and the user. It displays real-time market information, predicted Bitcoin price trends, and analytical insights in an understandable format for researchers, financial analysts, and investors.

The proposed architecture integrates streaming data technologies with deep learning prediction models, enabling continuous analysis of cryptocurrency market behavior. By combining data acquisition, processing, prediction, and visualization into a unified framework, the system provides an efficient platform for real-time Bitcoin price analysis and forecasting.

IV. HARDWARE DESCRIPTION

The implementation of the proposed Streaming Data Collection and Analysis system for Bitcoin price prediction requires a computing platform capable of handling real-time data streaming, machine learning model training, and result visualization. The hardware configuration mainly consists of a computer system with sufficient processing capability, memory, and storage to support data processing and deep learning operations. The main hardware components used in the system are described below.

Personal Computer / Processing System

The Personal Computer acts as the primary processing unit of the system. It is responsible for collecting streaming Bitcoin data, performing preprocessing operations, training the LSTM model, and generating prediction results. The computer executes the machine learning algorithms using programming environments such as Python and deep learning libraries. A multi-core processor improves the speed of data analysis and neural network computation. The processing system also manages real-time data streaming from cryptocurrency APIs and controls the overall operation of the prediction framework.

Central Processing Unit (CPU)

The Central Processing Unit performs all computational tasks involved in the system. It executes instructions related to data preprocessing, feature extraction, and LSTM model training. Since deep learning algorithms require significant computational resources, a high-performance processor such as Intel Core i5 or higher is recommended. The CPU ensures efficient execution of mathematical operations required for neural network processing and time-series prediction.

Memory (RAM)

Random Access Memory is required to temporarily store streaming data, processed datasets, and intermediate results during computation.

Cryptocurrency markets generate large volumes of real-time data, and sufficient memory capacity is essential for efficient data handling. A minimum of 8 GB RAM is recommended to support machine learning libraries, dataset loading, and model training without performance delays. Adequate memory availability improves system responsiveness and ensures smooth execution of the analysis process.

Storage Device

The storage device is used to store historical Bitcoin datasets, trained machine learning models, and analysis results. Solid State Drives (SSD) are preferred due to their faster read and write speeds compared to traditional hard disk drives. The stored historical data helps the system train the LSTM model and improve prediction accuracy. Storage is also used to maintain logs, prediction outputs, and graphical visualization results for further research and evaluation.

Internet Connectivity

Stable internet connectivity is essential for streaming real-time Bitcoin data from cryptocurrency exchange platforms. The system retrieves live market data through APIs provided by online exchanges. Continuous internet access ensures that the streaming module receives updated market information without interruption. Reliable connectivity is necessary for maintaining real-time analysis and accurate prediction of Bitcoin price trends.

V. RESULT AND DISCUSSION

The proposed Bitcoin streaming data analysis system was experimentally evaluated to study its data collection efficiency, prediction accuracy, and system reliability. The evaluation mainly focused on three aspects: real-time data streaming performance, LSTM prediction accuracy, and visualization of predicted market trends.

Real-Time Data Collection Performance

The streaming data collection module was tested by connecting the system to cryptocurrency exchange APIs that provide live Bitcoin market data. The system successfully collected continuous streams of data including Bitcoin price, trading volume, and timestamps. The streaming module demonstrated stable performance with minimal delay in receiving updated market information.

Multiple test sessions were conducted to verify the reliability of the streaming mechanism. The system was able to continuously collect and update the dataset without interruption. This ensured that the LSTM prediction model always had access to the most recent market data, improving the accuracy of future predictions.

LSTM Prediction Accuracy

The Long Short-Term Memory (LSTM) neural network model was trained using historical Bitcoin price data combined with newly streamed data. During the training phase, the model learned sequential patterns and long-term dependencies present in cryptocurrency price movements.

Experimental results showed that the LSTM model was capable of capturing complex market trends and producing predictions that closely followed actual Bitcoin price movements. The comparison between predicted and actual values demonstrated that the model successfully identified short-term fluctuations and long-term trends within the market.

The use of time-series learning allowed the LSTM network to remember important past patterns, which improved prediction accuracy compared to traditional machine learning models.

Data Visualization and Market Trend Analysis

The system generated graphical visualizations of both historical and predicted Bitcoin prices. Line graphs were used to represent price trends over time, allowing users to easily compare predicted values with actual market prices.

These visual representations provided clear insights into market behavior and helped identify price growth patterns, sudden fluctuations, and potential future trends. The visualization module enhanced the usability of the system by presenting complex analytical results in an understandable format.

System Stability and Processing Performance

During continuous operation, the system maintained stable performance without software failures or data

loss. The processing time required for data preprocessing, model prediction, and visualization remained within acceptable limits for real-time analysis.

Efficient memory management and optimized data handling techniques ensured that the system could process large datasets without significant performance degradation. The integration of streaming data technology with deep learning algorithms demonstrated reliable performance in real-time cryptocurrency market analysis.

Overall System Evaluation

The experimental evaluation confirms that the proposed system:

• Successfully collects real-time Bitcoin market data using streaming APIs.

• Efficiently preprocesses and organizes large volumes of financial data.

• Accurately predicts Bitcoin price trends using the LSTM deep learning model.
• Provides graphical visualization for better market analysis.
• Operates reliably for continuous real-time cryptocurrency monitoring.

VI. CONCLUSION AND FUTURE SCOPE

This paper presented the design and implementation of a streaming data collection and analysis system for Bitcoin price prediction using the Long Short-Term Memory (LSTM) deep learning algorithm. The proposed system integrates real-time cryptocurrency data acquisition, data preprocessing, deep learning-based prediction, and visualization to provide an effective framework for analyzing Bitcoin market trends.

The streaming data collection module successfully retrieves live Bitcoin market information from cryptocurrency exchange APIs, enabling continuous monitoring of price fluctuations and trading activity. The collected data is processed and organized into time-series sequences suitable for machine learning analysis. The LSTM neural network model effectively learns long-term dependencies in the sequential data and generates predictions of future Bitcoin price movements.

Experimental evaluation demonstrates that the LSTM model is capable of capturing complex patterns present in cryptocurrency markets and producing predictions that closely follow actual fprice trends. The integration of real-time streaming data with deep learning techniques enhances prediction reliability and ensures that the analysis reflects current market conditions.

The results confirm that the proposed system provides a practical and efficient solution for real-time Bitcoin data analysis and forecasting. The system can assist researchers, financial analysts, and cryptocurrency investors in understanding market behavior and making informed trading decisions. The implementation also demonstrates the potential of combining streaming data technologies with deep learning models for financial time-series analysis.

Future Scope

Although the proposed system demonstrates effective real-time Bitcoin data analysis and prediction using the LSTM algorithm, several improvements can be made to further enhance its accuracy, scalability, andf analytical capabilities.

Future work may include the integration of additional data sources such as social media sentiment analysis, news analytics, and blockchain transaction data to provide a more comprehensive understanding of cryptocurrency market behavior. Incorporating these external factors can improve prediction accuracy by capturing market sentiment and investor reactions.

Advanced deep learning models such as Gated Recurrent Units (GRU), Convolutional Neural Networks (CNN), or hybrid deep learning architectures can also be explored to improve forecasting performance. Ensemble learning techniques that combine multiple prediction models may further enhance system reliability.

The system can also be extended to support analysis of multiple cryptocurrencies such as Ethereum, Litecoin, and Ripple. Implementing cloud-based data processing and storage solutions can improve scalability and allow large-scale financial data analysis.

Additionally, the integration of automated trading strategies and real-time decision support systems can transform the proposed framework into a complete intelligent cryptocurrency trading platform. These improvements would enable the system to provide

more advanced financial analytics and predictive insights for modern digital asset markets.

REFERENCES

[1] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” Decentralized Business Review, 2008.

[2] J. McNally, J. Roche, and S. Caton, “Predicting the Price of Bitcoin Using Machine Learning,” in Proceedings of the IEEE International Conference on Data Science and Advanced Analytics, 2018.

[3] S. Patel, J. Shah, P. Thakkar, and K. Kotecha, “Predicting Stock and Cryptocurrency Market Using Machine Learning,” Expert Systems with Applications, vol. 130, pp. 54–63, 2019.

[4] I. E. Livieris, E. Pintelas, and P. Pintelas, “A CNN–LSTM Model for Cryptocurrency Price Prediction,” Neural Computing and Applications, Springer, 2020.

[5] D. Mallqui and R. Fernandes, “Predicting the Direction, Maximum, Minimum and Closing Prices of Daily Bitcoin Exchange Rate Using Machine Learning Techniques,” Applied Soft Computing, vol. 75, pp. 596–606, 2019.

[6] S. Lahmiri and S. Bekiros, “Cryptocurrency Forecasting with Deep Learning Chaotic Neural Networks,” Chaos, Solitons & Fractals, vol. 118, pp. 35–40, 2019.

[7] Y. Chen, Y. Li, and J. Wang, “Bitcoin Price Prediction Based on Deep Learning and LSTM Neural Networks,” Department of Computer Science, University of Technology, China, 2020.

[8] X. Huang, H. Zhang, and Y. Wang, “Cryptocurrency Price Prediction Using LSTM and Time Series Analysis,” IEEE International Conference on Big Data Analytics, 2021.

[9] A. Greaves and B. Au, “Using the Bitcoin Transaction Graph to Predict the Price of Bitcoin,” Department of Computer Science, Stanford University, USA, 2018.

[10] M. Mudassir, S. Bennbaia, D. Unal, and M. Hammoudeh, “Time-Series ForeLearning Models,” IEEE Access, vol. 8, pp. 123456–123470, 2020.

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

[12]S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.. Available: bitcoin.org

[13]A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd ed. Sebastopol, CA: O’Reilly Media, 2019.

[14]Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," IEEE Transactions on Neural Networks, vol. 5, no. 2, pp. 157–166, Mar. 1994.

[15]D. M. Q. Tran, M. Magdon-Ismail, et al., "Bitcoin Price Prediction Using LSTM, RNN and CNN," in Proc. IEEE Int. Conf. on Big Data, 2019, pp. 3425–3432.

[16]T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," European Journal of Operational Research, vol. 270, no. 2, pp. 654–669, 2018.

[17].Apache Kafka, "A distributed streaming platform," 2023.. Available: kafka.apache.org

[18]S.Siami-Namini, N. Tavakoli, and A. S. Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," in Proc. 17th IEEE Int. Conf. on Machine Learning and Applications, 2018, pp. 1394–1401.

[19]Coinbase, "Coinbase Pro API Reference," 2024.. Available: docs.cloud.coinbase.com

[20]F. Chollet, Deep Learning with Python, 2nd ed. Shelter Island, NY: Manning Publications, 2021.[Online]