iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.61584

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Analytical Study for Prjice Prediction of Bitcoin Using Machine Learning and Artificial Intelligence

Mrs. K. Kalaivani[1], Lagdhir Anjaliben Gopalbhai[2], S. Sandhiya[3], A. Suguna[4]
Assistant Professor of the Department Of Computer Application
Idhaya College of Arts and Science for Women, Pakkamudayanpet, Puducherry-605008

Abstract: ***Cryptocurrency transactions rely on encryption algorithms for security and decentralization. However, traditional machine learning algorithms often struggle with the dynamic cryptocurrency market, leading to inaccuracies. To address this, integrating Integrating sentiment analysis of Twitter data alongside Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) models enriches the analysis of cryptocurrency market dynamics. Twitter provides dataset from kaggle open source and the sentiment expressed by users on cryptocurrencies, enabling the prediction of tweet sentiment (positive, negative, or neutral) through Natural Language Processing (NLP). This integration enhances understanding of market sentiment's impact on cryptocurrency prices: positive sentiment may drive prices up, negative sentiment may lead to declines, and neutral sentiment may indicate stability. By analyzing sentiment alongside historical trends and emerging patterns, users gain a holistic view of cryptocurrency markets. This approach aids decision-making, improving transaction accuracy and efficiency. Ultimately, the combination of sentiment analysis with BiLSTM and BiGRU models advances the comprehension of cryptocurrency market dynamics, enhancing user insights and facilitating informed decisions in the volatile cryptocurrency ecosystem.***

Keywords: cryptocurrency, digital payment, encryption algorithms, secure transactions, decentralized, Bidirectional Long Short-Term Memory (BiLSTM) models,transactional activities, market dynamics, past dependencies, future dependencies, sequential data.

I. INTRODUCTION

Cryptocurrency comes under many names. You have probably read about some of the most popular types of cryptocurrencies such as Bitcoin, Litecoin, and Ethereum. Cryptocurrencies are increasingly popular alternatives for online payments. Before converting real dollars, euros, pounds, or other traditional currencies into Ƀ (the symbol for Bitcoin, the most popular cryptocurrency), you should understand what cryptocurrencies are, what the risks are in using cryptocurrencies, and how to protect your investment. A cryptocurrency is a digital currency, which is an alternative form of payment created using encryption algorithms. The use of encryption technologies means that cryptocurrencies function both as a currency and as a virtual accounting system. To use cryptocurrencies,you need a cryptocurrency wallet. These wallets can be software that is a cloud-basedservice or is stored on your computer or on your mobile device. The wallets are the tool through which you store your encryption keys that confirm your identity and link to your cryptocurrency. Cryptocurrencies are still relatively new, andthe market for these digital currencies is very volatile. Since cryptocurrencies don't need banks or any other third party to regulate them; they tend to be uninsured and are hardto convert into a form of tangible currency (such as US dollars or euros.) In addition, since cryptocurrencies are technology-basedintangible assets, they can be hacked like anyother intangible technology asset. Finally,since you store your cryptocurrencies in a digital wallet, if you lose your wallet (oraccess to it or to wallet backups), you have lost your entire cryptocurrency investment. Look before you leap! Before investing in a cryptocurrency, be sure you understand howit works, where it can be used, and how to exchange it. Read the webpages for the currency itself (such as Ethereum, Bitcoin orLitecoin) so that you fully understand how it works, and read independent articles on the cryptocurrencies you are considering as well.

II. LITERATURE SURVEY

1) *Cryptocurrency Price Prediction using Machine Learning Algorithm[1]* In the realm of cryptocurrency trading, accurate predictions are paramount. Leveraging machine learning algorithms like Linear Regression, Random Forest Regressor, Gradient Boosting Regressor, and XGBoost, our study aimed to analyze Bitcoin, XRP, Ethereum, and Stellar markets. With validation accuracy rates reaching 95–97%, these models effectively forecast daily price behaviors. This success empowers investors and traders with insights to navigate market volatility confidently. By understanding intricate market dynamics, informed decisions become feasible. The integration of advanced machine learning algorithms not only enhances prediction precision but also addresses the rapid and dynamic fluctuations inherent in cryptocurrency environments.

©IJRASET: All Rights are Reserved

2) *Cryptocurrency Price Prediction Based on Long-Term and Short-Term Integrated Learning [2]* This paper introduces a cryptocurrency price expectation model based on Support Vector Regression (SVR) with a unique approach called long-term and short-term integrated learning. By analyzing a vast dataset of historical cryptocurrency prices, the model aims to improve price prediction accuracy by considering both extended trends and immediate market dynamics. SVR, known for handling nonlinear relationships in data, forms the foundation of this model. By combining long-term factors like historical price patterns and market sentiment with short-term elements such as intraday trading patterns and real-time market volatility, the model gains a comprehensive understanding of the cryptocurrency market. This integration of diverse temporal components enhances the model's adaptability and precision in predicting cryptocurrency prices, facilitating informed decision-making in this rapidly evolving digital economy.
3) *Cryptocurrency Price Prediction Using Deep Learning [3]* This project aims to utilize deep learning algorithms, specifically Long Short-Term Memory (LSTM), to forecast Bitcoin prices, acknowledging its evolution into an investment asset and the necessity for accurate predictions in the volatile cryptocurrency market. In the initial phase, historical cryptocurrency data trains the LSTM model to discern patterns and trends in Bitcoin's price movements, leveraging LSTM's ability to capture long-term dependencies. In the second phase, a frontend application is developed for practical use, enabling users to access Bitcoin price predictions for specific timeframes. This user-friendly interface enhances accessibility, empowering investors and enthusiasts to make informed decisions based on tailored forecasts. Through this approach, the project bridges the gap between deep learning research and real-world applications, offering valuable insights into Bitcoin price dynamics and facilitating informed decision-making in the cryptocurrency market.
4) *Bitcoin Price Prediction Using Machine Learning and Deep Learning Algorithms [4]* The paper delves into the evolution of cryptocurrencies from transactional tools to investment assets and highlights the need for accurate price predictions due to their inherent volatility. Focusing on Bitcoin, the study evaluates the performance of various machine learning algorithms in predicting its closing prices, crucial for informed investment decisions. Metrics such as accuracy and mean absolute error are employed for assessment. The research identifies the Auto Regressive Integrated Moving Average (ARIMA) model as the most effective, showcasing the least mean absolute error. ARIMA's success stems from its adeptness in capturing temporal patterns inherent in Bitcoin price data, emphasizing the importance of selecting models aligned with cryptocurrency characteristics. Furthermore, the study emphasizes the significance of quality training data and dataset volume for successful predictions. It validates ARIMA's predictive accuracy by comparing predicted Bitcoin values with actual ones over a four-month analysis period. Overall, the research underscores the importance of appropriate model selection and comprehensive historical data in accurately predicting cryptocurrency prices.
5) *Cryptocurrency Price Prediction using Time Series Forecasting (ARIMA) [5]* Cryptocurrency's prominence in finance spurred a study using machine learning, including Linear Regression, Random Forest, Gradient Boosting, and XGBoost, to analyze Bitcoin, XRP, Ethereum, and Stellar. This diverse approach unveiled intricate market patterns, yielding impressive validation accuracy rates of 95–97%. Accurate predictions empower investors to navigate volatile markets confidently. The success of these models highlights their efficacy in deciphering cryptocurrency market dynamics, aiding informed decision-making. This integration of advanced machine learning algorithms enhances financial precision amid rapid market fluctuations, offering significant implications for traders and investors seeking to navigate the cryptocurrency landscape with confidence.

III. IMPLEMENTATION

A. Data Extraction

Data is extracted from the provided API endpoint, specifically from the URL format: `https://min-api.cryptocompare.com/data/v2/histo'+frequency+'?'+cryptoHeader`, where `frequency` represents the time interval for historical data (e.g., hourly, daily), and `cryptoHeader` includes parameters such as the cryptocurrency symbol, currency pair, and any other relevant details. Once retrieved, the data is stored as a CSV (Comma-Separated Values) file, a common format for tabular data. This file likely contains columns representing various attributes such as timestamp, cryptocurrency price, volume, and other market-related metrics.

The extracted data serves as the foundation for analysis and prediction in cryptocurrency trading. It enables researchers and analysts to study market behaviors, trends, and fluctuations over time, providing valuable insights for decision-making processes. Overall, this process involves retrieving historical cryptocurrency data from the specified API endpoint, storing it in a CSV file format for further analysis, and leveraging this data to gain insights into cryptocurrency market dynamics.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

B. Pre-processing

Pre-processing of cryptocurrency data involves several crucial steps to ensure its suitability for analysis and modeling. Initially, the data undergoes cleaning, where any inconsistencies, outliers, or missing values are addressed to enhance its quality and reliability. Subsequently, feature engineering techniques may be applied to extract relevant information or create new variables that better capture underlying patterns in the data. This can include transforming timestamps into meaningful features or calculating additional indicators to provide deeper insights into market behavior. Normalization or standardization may also be performed to ensure that all features are on a similar scale, preventing biases in the analysis. Furthermore, the data may be segmented into training and testing sets, with attention to preserving temporal order, especially in the case of time-series data like cryptocurrency prices. By effectively pre-processing the data, analysts can ensure that it is ready for accurate and meaningful analysis, laying the groundwork for informed decision-making in cryptocurrency trading and investment.

C. Feature Extraction

Feature extraction in the context of cryptocurrency data analysis involves identifying and deriving relevant attributes or features from raw data to facilitate effective analysis and modeling. This process aims to capture essential information that can contribute to understanding market dynamics and predicting price movements. In cryptocurrency data, features may include various metrics such as price volatility, trading volume, market sentiment indicators, and technical analysis indicators like moving averages or relative strength index (RSI). Feature extraction techniques may involve mathematical transformations, statistical calculations, or domain-specific knowledge to create meaningful representations of the underlying data. By selecting and extracting pertinent features, analysts can uncover valuable insights and patterns within the cryptocurrency market, aiding in decision-making processes for traders and investors. Effective feature extraction is essential for optimizing the performance of machine learning models and enhancing the accuracy of cryptocurrency price predictions.

D. Model Creation

The creation of a model using Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) represents a sophisticated approach to cryptocurrency price prediction. These models leverage their unique architectures to capture both past and future dependencies in sequential cryptocurrency data. BiLSTM, known for its ability to retain long-term memory, can effectively analyze historical trends and patterns in cryptocurrency prices. On the other hand, BiGRU, with its efficient processing of sequential data, excels in capturing short-term fluctuations and emerging trends. By integrating both BiLSTM and BiGRU layers in the model, it can simultaneously analyze historical data while considering real-time market dynamics, providing a comprehensive understanding of cryptocurrency market behavior. This bidirectional approach enables the model to capture complex patterns and relationships in the data, ultimately enhancing the accuracy of price predictions. The utilization of BiLSTM and BiGRU in cryptocurrency modeling signifies a cutting-edge methodology that addresses the challenges posed by the dynamic and volatile nature of cryptocurrency markets, offering valuable insights for traders and investors.

1) BI-LSTM

Bidirectional Long Short-Term Memory (BiLSTM) is a variant of the Long Short-Term Memory (LSTM) neural network architecture, designed to model sequential data while considering both past and future information. In traditional LSTM networks, information flows only in one direction, typically from past to future. However, BiLSTM addresses this limitation by incorporating two LSTM layers: one processing the input sequence in the forward direction and another processing it in the backward direction.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

Candidate Value:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Output Gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

Cell State Update:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Hidden State Update:

$$h_t = o_t * \tanh(C_t)$$

This bidirectional processing enables the model to capture dependencies from both past and future contexts simultaneously, enhancing its ability to understand and predict sequential data. At each time step, the forward LSTM cell processes the input sequence from the beginning to the end, while the backward LSTM cell processes it from the end to the beginning.

- f_t, i_t, o_t are the forget, input, and output gate vectors, respectively.
- $\tilde{C}_t$ is the candidate value vector.
- C_t is the cell state vector.
- h_t is the hidden state vector.
- W_f, W_i, W_o, W_C are weight matrices, and b_f, b_i, b_o, b_C are bias vectors.
- $[h_{t-1}, x_t]$ represents the concatenation of the previous hidden state h_{t-1} and the current input x_t.
- σ represents the sigmoid activation function, and $\tanh$ represents the hyperbolic tangent activation function.

The outputs from both directions are then concatenated, allowing the model to access information from both past and future contexts when making predictions. This bidirectional approach enables BiLSTM to capture complex patterns and long-range dependencies in sequential data, making it particularly well-suited for tasks such as speech recognition, language modeling, and time series prediction, including forecasting cryptocurrency prices.

2) BI-GRU

Bidirectional Gated Recurrent Unit (BiGRU) is a variation of the Gated Recurrent Unit (GRU) architecture, designed to capture bidirectional dependencies in sequential data. Similar to BiLSTM, BiGRU processes input sequences in both forward and backward directions simultaneously, enabling the model to leverage information from past and future contexts when making predictions. The key components of a BiGRU cell include update and reset gates, which control the flow of information within the cell, and a candidate value that represents the new information to be added to the cell state.

Update Gate:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z)$$

Reset Gate:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r)$$

Candidate Value:

$$\tilde{h}_t = \tanh(W_h \cdot [r_t * h_{t-1}, x_t] + b_h)$$

Hidden State Update:

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

During each time step, the forward and backward GRU cells process the input sequence independently, capturing relevant patterns and dependencies. The outputs from both directions are then combined, typically through concatenation, allowing the model to access information from both past and future contexts. This bidirectional approach enhances the model's ability to understand and predict sequential data, making BiGRU suitable for tasks such as natural language processing, time series analysis, and sentiment analysis.

- z_t and r_t are the update and reset gate vectors, respectively.
- $\tilde{h}_t$ is the candidate value vector.
- h_t is the hidden state vector.
- W_z, W_r, W_h are weight matrices, and b_z, b_r, b_h are bias vectors.
- $[h_{t-1}, x_t]$ represents the concatenation of the previous hidden state h_{t-1} and the current input x_t.
- σ represents the sigmoid activation function, and $\tanh$ represents the hyperbolic tangent activation function.

Overall, BiGRU offers a powerful framework for capturing bidirectional dependencies in sequential data, contributing to its effectiveness in various machine learning applications.

E. Processing the Twitter Data

To process Twitter data for cryptocurrency analysis, first, collect tweets using the Twitter kaggle dataset based on relevant keywords or hashtags. Then, clean the data by removing noise like URLs and hashtags and tokenize the text. Perform sentiment analysis to classify tweets as positive, negative, or neutral using NLP techniques. Extract features such as word frequency and sentiment scores. Encode the data into a suitable format for machine learning models. Integrate the processed Twitter data with historical cryptocurrency data and feed it into Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) models for analysis. Evaluate the model's performance and iterate to improve accuracy and efficiency. This approach enables a comprehensive understanding of market sentiment and enhances decision-making in the cryptocurrency ecosystem.

IV. RESULT AND DISCUSSION

In the proposed system, sentiment analysis of Twitter data sourced from Kaggle's open dataset is integrated alongside Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) models to enhance the analysis of cryptocurrency market dynamics. Leveraging Natural Language Processing (NLP) techniques, the sentiment expressed by users on cryptocurrencies in tweets is predicted as positive, negative, or neutral. This integration provides valuable insights into market sentiment and its impact on cryptocurrency prices: positive sentiment may drive prices up, negative sentiment may lead to declines, and neutral sentiment may indicate market stability. By analyzing sentiment alongside historical trends and emerging patterns, users obtain a comprehensive view of cryptocurrency markets, aiding decision-making processes. This holistic approach improves transaction accuracy and efficiency, advancing the comprehension of cryptocurrency market dynamics and empowering users to make informed decisions in the volatile cryptocurrency ecosystem.

A. Train and Validation Loss



The training loss represents the error or discrepancy between the predicted cryptocurrency prices generated by the models during the training phase and the actual prices observed in the training dataset. It is calculated using a loss function, such as Mean Squared Error (MSE) or Mean Absolute Error (MAE), which quantifies the difference between predicted and actual values. The formula for calculating training loss with MSE is:

$$\text{Training Loss} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Where:

- n is the number of samples in the training dataset.
- y_i is the actual cryptocurrency price for the i^{th} sample.
- $\hat{y}_i$ is the predicted cryptocurrency price for the i^{th} sample.

The validation loss measures the error between the predicted and actual cryptocurrency prices on a separate validation dataset, which the model has not been exposed to during training. It serves as an estimate of the model's performance on unseen data and helps detect overfitting. The formula for calculating validation loss is the same as for training loss:

$$\text{Validation Loss} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Where:

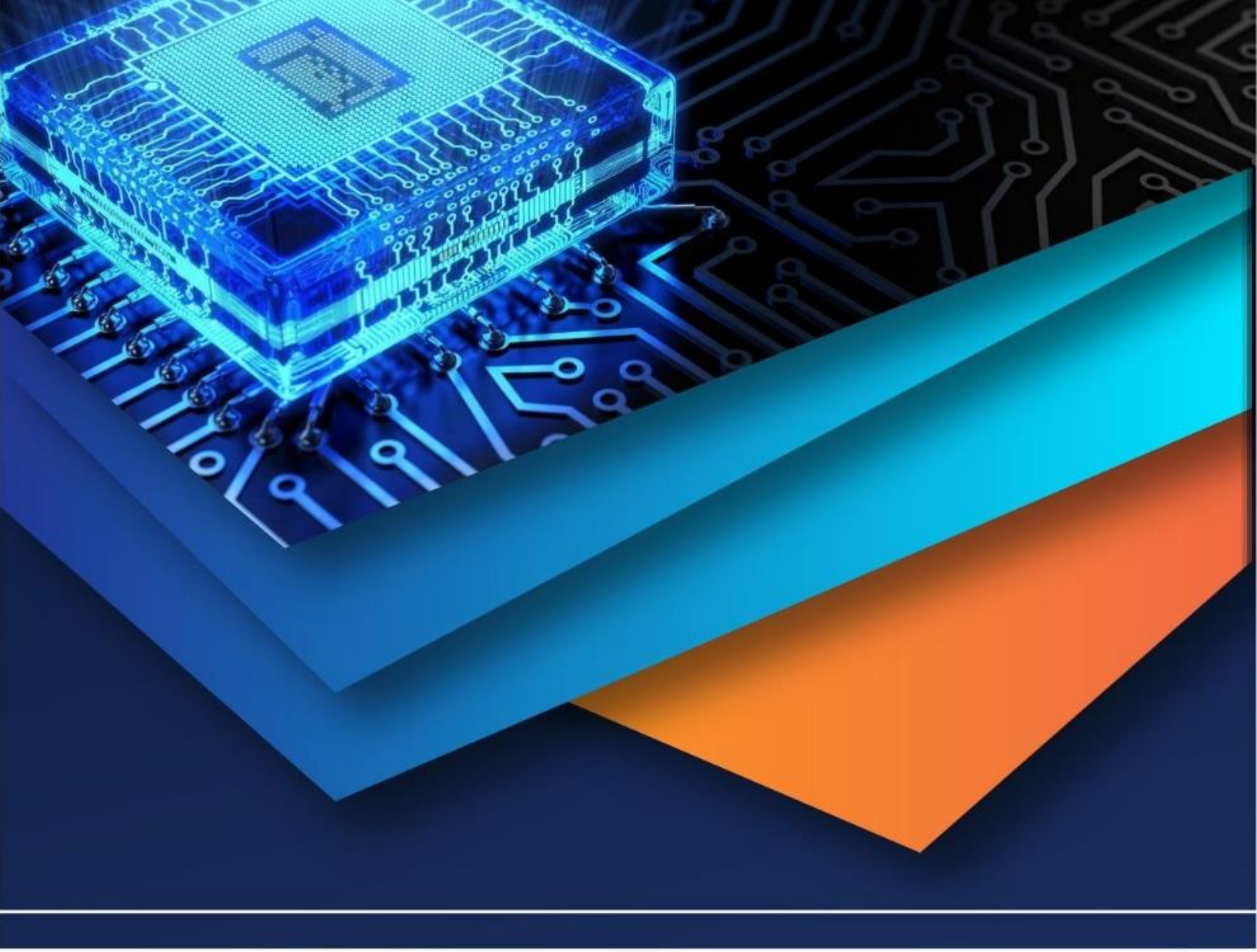
- n is the number of samples in the validation dataset.
- y_i is the actual cryptocurrency price for the i^{th} sample in the validation dataset.
- $\hat{y}_i$ is the predicted cryptocurrency price for the i^{th} sample in the validation dataset.

V. CONCLUSION

In conclusion, the proposed system leverages sentiment analysis of Twitter data alongside Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) models to enhance the analysis of cryptocurrency market dynamics. By incorporating real-time sentiment expressed by users on cryptocurrencies, the system provides valuable insights into market sentiment's impact on cryptocurrency prices. Through Natural Language Processing (NLP) techniques, tweets are classified as positive, negative, or neutral, enriching the understanding of how sentiment influences cryptocurrency prices. The integration of sentiment analysis alongside historical trends and emerging patterns enables users to gain a holistic view of cryptocurrency markets, aiding decision-making processes. This holistic approach improves transaction accuracy and efficiency, ultimately advancing the comprehension of cryptocurrency market dynamics and empowering users to make informed decisions in the volatile cryptocurrency ecosystem. Future work could focus on refining the sentiment analysis model to better capture nuances in cryptocurrency-related tweets and improve accuracy. Additionally, exploring advanced deep learning architectures or ensemble techniques for predicting cryptocurrency prices could enhance the system's performance. Integration of other external data sources, such as news articles or forum discussions, could further enrich the analysis and provide deeper insights into market dynamics. Finally, developing a user-friendly interface for accessing and visualizing the analysis results could enhance usability and adoption of the system among cryptocurrency investors and traders.

REFERENCES

[1] Monisha Mittal; G. Geetha, "Predicting Bitcoin Price using Machine Learning" in IEEE 2022.
[2] Akhilesh Kumar Singh; Manish Raj, "A study of Analytics and Exploration of Bitcoin Challenges and Blockchain" in 2022 IEEE.
[3] Grace. LK. Joshila; Asha. P; D. Usha Nandini; G. Kalaiarasi, "Price Prediction of Bitcoin" in IEEE 2021.
[4] Raj Gaurang Tiwari; Ambuj Kumar Agarwal; Rajesh Kumar Kaushal; Naveen Kumar, “Prophetic Analysis of Bitcoin price using Machine Learning Approaches”, in IEEE 2021.
[5] Wenhan Hou; Bo Cui; Ru Li, “A Survey on Blockchain Data Analysis” in IEEE 2021.
[6] Mayukh Samaddar*1, Rishiraj Roy*1, Sayantani De*1 and Raja Karmakar# ,“ A Comparative Study of Different Machine Learning Algorithms on Bitcoin Value Prediction” in IEEE 2021.
[7] Brian Vockathaler, “ The Bitcoin Boom: An In Depth Analysis of The Price Of Bitcoins”, Thesis, University Of Ottawa, Ontario, Canada, June 2017
[8] Huisu Jang and Jaewook Lee, “An Empirical Study on Modelling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Blockchain Information,”in IEEE Early Access Articles, 2017.
[9] Issac Madan, Shaurya Saluja, Aojia Zhao,Automated Bitcoin Trading via Machine Learning Algorithms, Department of Computer Science , Stanford University, Stanford , 2015.
[10] F. Andrade de Oliveira, L. Enrique ZÃ¡rate and M. de Azevedo Reis; C. Neri Nobre, “The use of artificial neural networks in the analysis and prediction of stock prices,” in IEEE International Conference on Systems, Man, and Cybernetics, 2011.
[11] M. Daniela and A. BUTOI, “Data mining on Romanian stock market using neural networks for price prediction”. informatica Economica, 2013



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)