iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68470

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Cryptocurrency Price Prediction using Machine Learning

Mr. R.Prapulla Kumar[1], B. Krishnaveni[2], R. Manvitha[3], T. Sruthi[4], T. Padma[5]
[1]*Assistant Professor, Department of CSE, PBR VITS Kavali, Nellore–524201*
[2]*TeamLeader, Student PBR VITS Kavali, Nellore–524201*
[3, 4, 5]*Student PBRVITS Kavali, Nellore–524201*

Abstract: Cryptocurrency markets exhibit high volatility, making accurate price prediction a challenging task. This project aims to develop a cryptocurrency price prediction model using linear regression. The model is trained on historical price data, considering key features such as closing prices, trading volume, and market trends. The implementation is built with Streamlit, allowing for an interactive and user-friendly interface where users can input parameters and visualize predictions dynamically.
Keywords: Cryptocurrency, Machine Learning, Linear Regression, Price Prediction, Data Analytics,Streamlit, Trading Insights, Financial Forecasting,Interactive Visualization,Historical Data.

I. INTRODUCTION

Cryptocurrencies have gained significant popularity in recent years due to their decentralized nature and high return potential. However, their prices are highly volatile, making it challenging for traders and investors to make informed decisions. Accurate price prediction models can provide valuable insights for better financial planning and risk management.This project focuses on building a cryptocurrency price prediction system using linear regression, a widely used statistical method for forecasting numerical values. Historical price data, trading volume, and market trends are analyzed to train the model and improve prediction accuracy. The project is implemented using Streamlit, enabling users to interactively visualize predictions and explore historical trends through an intuitive web-based interface.

A. Objectives of the Study:

1) Develop a cryptocurrency price prediction model using linear regression.
2) Analyze and preprocess historical cryptocurrency price data.
3) Evaluate the model's performance using appropriate metrics.
4) Build an interactive web application using Streamlit.
5) Provide data-driven insights for informed investment decisions.

II. EXISTING SYSTEM

A. Existing Cryptocurrency Price Prediction Methods

Traditional methods for cryptocurrency price prediction primarily relied on statistical models, technical analysis, and fundamental analysis. Statistical models such as Moving Averages (MA, SMA, EMA) were commonly used to identify trends based on past prices, while Autoregressive Integrated Moving Average (ARIMA) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) were applied for time series forecasting and volatility modeling, respectively.However, these methods often struggled with highly volatile assets like cryptocurrencies.

Technical analysis (TA) played a significant role in traditional prediction methods, where traders used indicators such as Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands to identify market trends and potential buy/sell signals. Additionally, support and resistance levels were used to determine price reversal points. While TA provided useful insights, it was often subjective and not always reliable in unpredictable markets.

Fundamental analysis (FA) focused on external factors affecting cryptocurrency prices, including market sentiment, regulations, adoption rates, and on-chain metrics like Bitcoin's hash rate, active addresses, and transaction volume. Unlike statistical models, FA aimed to assess the intrinsic value of a cryptocurrency rather than relying purely on historical price movements.

Although these traditional methods provided some insights, they had limitations in handling complex, non-linear relationships in cryptocurrency price movements.

This led to the adoption of machine learning models, such as Linear Regression, XGBoost, and Long Short-Term Memory (LSTM) networks, which can analyze large datasets and capture intricate patterns more effectively. Your project, which applies Linear Regression with Streamlit, improves upon these traditional methods by leveraging historical price data and statistical relationships to make more data-driven predictions.

B. LimitationsofTraditionalApproaches

One major drawback of statistical modelslikeMoving Averages (SMA, EMA)andARIMA is that they assume past price trends will continue, which is not always true in the unpredictable cryptocurrency market. These models struggle with sudden price spikes caused by news, regulations, or market sentiment. Additionally, ARIMA and GARCH are effective only for short-term predictions and fail to capture long-term price movements.

Technical analysis (TA) also has its limitations. Indicators like RSI, MACD, and Bollinger Bands depend heavily on historical price patterns and do not account for external factors such as government regulations, whale movements, or security breaches. Moreover, TA is often subjective, as different traders may interpret the same indicators differently, leading to inconsistent predictions.

Fundamental analysis (FA), while valuable in traditional finance, is challenging in the crypto space due to the lack of standardized financial reports and regulatory frameworks. Many cryptocurrencies lack clear intrinsic value, and speculative trading often drives prices more than real-world adoption or project fundamentals. Additionally, FA does not provide precise price predictions but rather gives a general outlook on market trends.

Overall, traditional methods struggle with non-linearity, sudden market shifts, and lack of adaptability to new patterns. This is why machine learning models, such as Linear Regression, XGBoost, and LSTMs, are now preferred, as they can process large datasets and identify complex relationships more effectively. However, even ML models have challenges, such as data quality issues and overfitting.

III. RELATED WORK

Several studies and projects have explored cryptocurrency price prediction using various methodologies. Traditional approaches, such as time series forecasting models like ARIMA (AutoRegressive Integrated Moving Average)andGARCH (Generalized Autoregressive Conditional Heteroskedasticity)**,** have been used to analyze historical price trends. For example, a study by Patel et al. (2018) applied ARIMA for Bitcoin price prediction but found that the model struggled with high volatility.

Machine learning techniques, including Support Vector Machines (SVM)**,** Random Forests, andArtificial Neural Networks (ANNs), have been explored to improve prediction accuracy.

Early approaches relied on statistical models like ARIMA and GARCH, which were effective for short-term forecasting but struggled with the highly volatile nature of cryptocurrencies. Researchers found that while these models could capture linear price trends, they failed to account for sudden market fluctuations caused by external factors such as regulations, hacks, or investor sentiment.

With advancements in machine learning, researchers started using regression models such as Linear Regression (LR)**,** Support Vector Regression (SVR)**,** andRandom Forest Regression**.**

A study comparing these models found that LR performed well for short-term predictions, but it was less accurate for long-term forecasting due to its inability to capture non-linear relationships in price movements.

Additionally, some projects have combined technical indicators (e.g., RSI, MACD) with machine learning models to enhance predictive accuracy. Hybrid approaches, such as combining LSTM with ARIMA **or** XGBoost, have also been proposed to leverage the strengths of both traditional and AI-based methods.

IV. PROPOSED SYSTEM

A. Objective

The proposed system aims to develop a cryptocurrency price prediction model using linear regression and deploy it through an interactive Streamlit application. The goal is to provide a simple, interpretable, and user-friendly tool for traders and analysts.

By leveraging machine learning techniques, the model will provide transparent and interpretable forecasts, making it accessible to both technical and non-technical users. Additionally, the system will be designed to optimize computational efficiency, ensuring quick and reliable predictions without the need for extensive resources.

Specific Objectives:

1) *Accurate Price Prediction:*
- Utilize Linear Regression to analyze historical cryptocurrency prices and predict future price movements based on past trends.

2) *User-Friendly Interface with Streamlit:*
- Develop an interactive web application that allows users to input historical data, visualize price trends, and view predictions in real-time.

3) *Data Processing and Feature Engineering:*
- Clean and preprocess cryptocurrency price data, incorporating key indicators such as moving averages, trading volume, and volatility to enhance model performance.

4) *Performance Evaluation:*
- Assess the accuracy of the model using metrics such as Root Mean Squared Error (RMSE) and R^2 Score, ensuring reliable predictions.

5) *Scalability and Future Enhancements:*
- Design the system in a way that allows easy integration of advanced models (e.g., LSTMs, XGBoost) or additional features like sentiment analysis for improved accuracy.

B. Dataset Preparation and Model Training:

Data preprocessing techniques, including feature scaling, outlier removal, and handling missing values, will be applied to improve model performance. A linear regression algorithm will be used to establish relationships between selected features and future cryptocurrency prices. The model will be evaluated using metrics such as Mean Squared Error (MSE) and R^2 score to ensure accuracy.

Data preprocessing is crucial to ensure the accuracy and reliability of the prediction model. The steps involved include:

1) *Data Collection*
- Fetch historical cryptocurrency price data using APIs (e.g., Binance, CoinGecko) or load from CSV files.
- Select key features like Open, High, Low, Close (OHLC) prices, trading volume, and timestamps**.**

2) *Handling Missing Data*
- Check for missing values using df.isnull().sum().
- Fill missing values using interpolation**,** mean/median imputation, or drop them if necessary.

3) *c. Feature Engineering*
- Create new features such as moving averages (SMA, EMA), price volatility, and trading volume changes.
- Convert timestamps into meaningful features like day of the week, month, or trend indicators**.**

4) *d. Data Normalization*
- Normalize numerical features using MinMaxScalerorStandardScalerto improve model performance.

5) *Splitting Data*
- Split the dataset into training (80%)andtesting (20%) sets using train_test_split() from scikit-learn.

V. IMPLEMENTATION

A. System Setup and ModelTraining:

The system setup begins with the installation of essential libraries and frameworks required for data processing, model training, and web application development. Key dependencies include Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, and Streamlit. The environment is configured using Jupyter Notebook or any Python IDE, ensuring smooth data manipulation and model experimentation.

Scikit-learn is utilized for machine learning tasks, providing tools for data preprocessing, model selection, and evaluation. Matplotlib and Seaborn are employed for data visualization, enabling insightful graphical representations of historical price trends, correlations, and model performance metrics. Additionally, Streamlit is integrated to develop an interactive web-based user interface for real-time cryptocurrency price predictions.

To ensure a seamless development experience, the environment is configured using Jupyter Notebook, VS Code, or PyCharm, allowing efficient data exploration, model experimentation, and debugging.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

The system also supports API integration for real-time data fetching from sources such as Binance, CoinGecko, or Yahoo Finance, ensuring access to up-to-date cryptocurrency market trends.

B. PerformanceEvaluation

The performance of the cryptocurrency price prediction model is evaluated using standard regression metrics to ensure accuracy and reliability. Key metrics include Mean Squared Error (MSE), which measures the average squared difference between predicted and actual values, and Root Mean Squared Error (RMSE), which provides a more interpretable error measurement by taking the square root of MSE.

To validate the model, the dataset is split into training and testing sets (typically 80-20%), ensuring that the model generalizes well to unseen data. Cross-validation techniques may also be applied to further improve robustness.

Hyperparameter tuning methods like Grid Search and Randomized Search are employed to optimize model performance. The final trained model is integrated into a Streamlit-based web application, allowing users to visualize price predictions, assess accuracy, and explore potential improvements. Future enhancements may involve ensemble learning techniques **or** deep learning architectures (LSTMs, XGBoost) for better capturing complex cryptocurrency market patterns.

Additionally, the R^2 score (coefficient of determination**)** is used to assess how well the model explains the variance in the data, with values closer to 1 indicating better performance.

VI. FUTURE SCOPE

A. Advanced Machine Learning Model:

Future enhancements can include integrating LSTMs (Long Short-Term Memory networks), XGBoost, and deep learning models to improve prediction accuracy by capturing complex market patterns. These models can handle sequential dependencies in financial data, leading to more precise forecasts.

B. Real-Time Data Integration:

The system can be upgraded to fetch real-time cryptocurrency price data from exchanges, allowing users to make dynamic and up-to-date predictions. Live data streaming will improve the model's responsiveness to market fluctuations.

VII. CONCLUSION

The proposed cryptocurrency price prediction system provides an efficient and user-friendly approach to forecasting price trends using linear regression and an interactive Streamlit application. By leveraging historical price data, feature engineering, and machine learning techniques, the system offers a transparent and interpretable model for traders, analysts, and investors. The integration of data preprocessing techniques, performance evaluation, and visualization tools ensures that predictions are accurate and easy to understand.

VIII. FIGURES

Fig.1.Price Prediction.

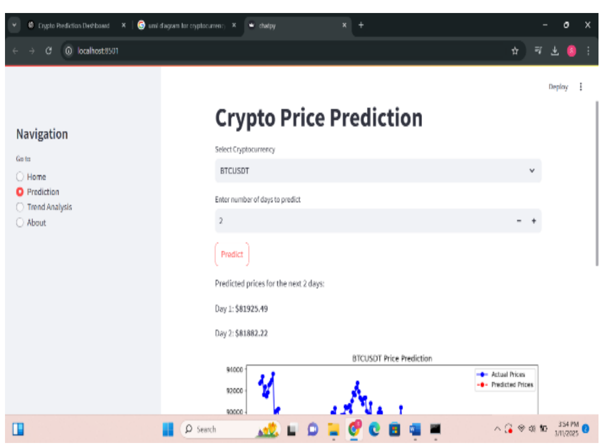


Fig.2.Predicted Graph.

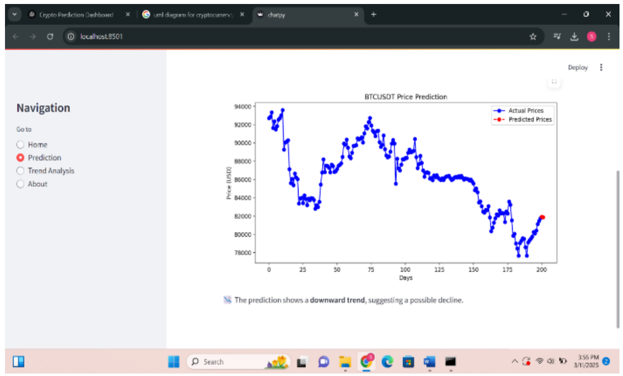


Fig.3.Trend Analysis.

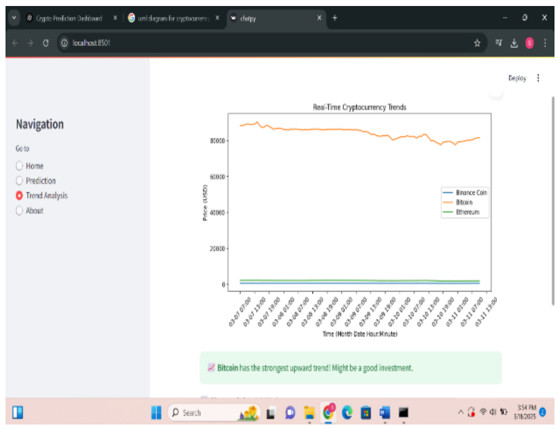
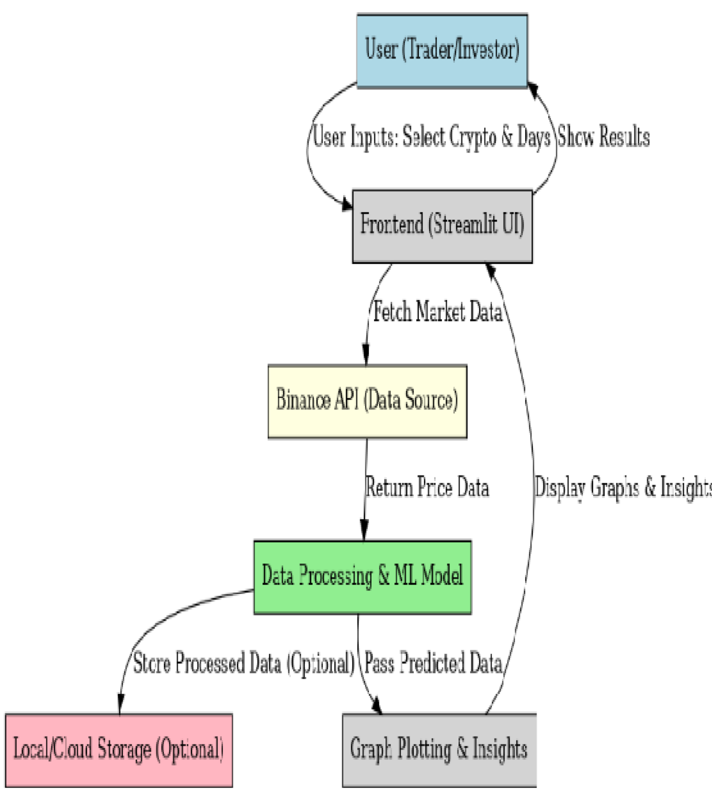


Fig.4.Architecture.



International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Fig.5.UML Diagram.

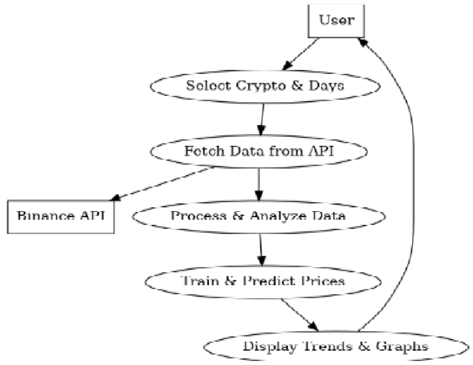
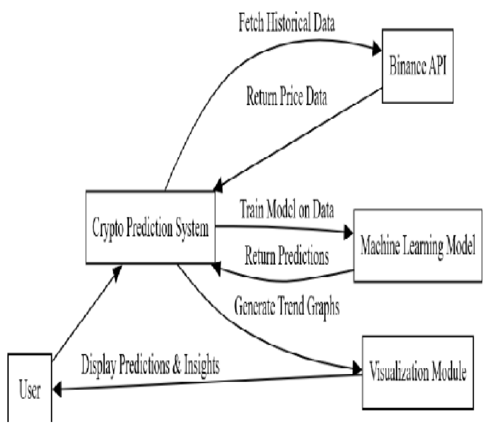
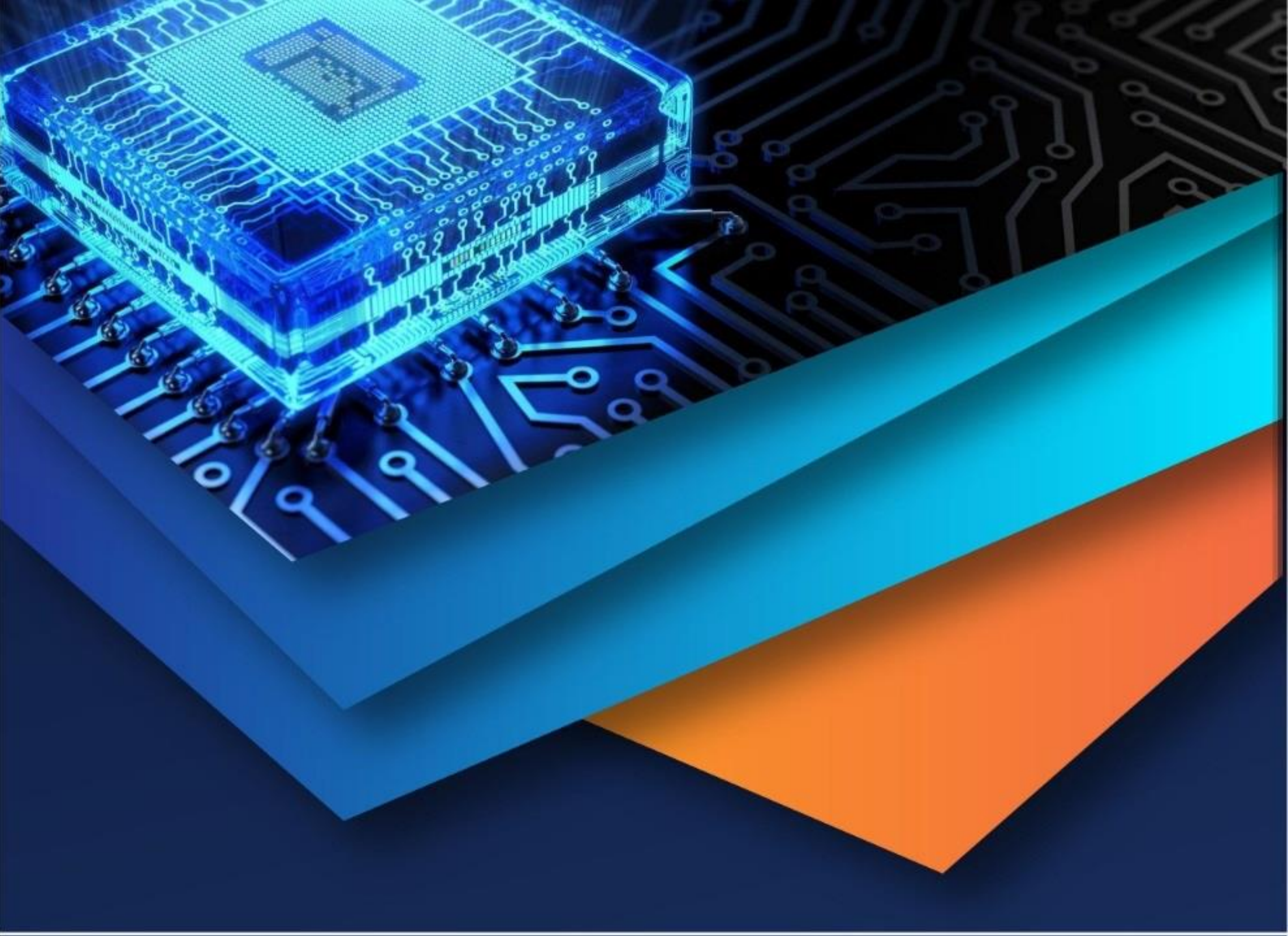


Fig.6.Sequence Diagram.



REFERENCES

[1] Box, G. E. P., & Jenkins, G. M. (1976).Time Series Analysis: Forecasting and Control. Holden-Day.
[2] Murphy, J. J. (1999).Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. New York Institute of Finance.
[3] Tsay, R. S. (2005).Analysis of Financial Time Series. Wiley.
[4] Hunter, J. D. (2007).Matplotlib: A 2D Graphics Environment.Computing in Science & Engineering, 9(3), 90-95. Retrieved from https://matplotlib.org/
[5] Nakamoto, S. (2008).Bitcoin: A Peer-to-Peer Electronic Cash System. Retrieved from https://bitcoin.org/bitcoin.pdf.
[6] Chong, Y. Y., & Ng, T. H. (2008).Technical analysis and the London stock exchange: Testing the MACD and RSI rules using the FT30.Applied Economics Letters, 15(14), 1111-1114.
[7] Hastie, T., Tibshirani, R., & Friedman, J. (2009).The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
[8] Chan, E. (2009).Quantitative Trading: How to Build Your Own Algorithmic Trading Business. Wiley.
[9] McKinney, W. (2010).Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly Media.
[10] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & others. (2011).Scikit-Learn: Machine Learning in Python.Journal of Machine Learning Research, 12, 2825–2830. Retrieved from https://scikit-learn.org/.
[11] Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015).Performance Analysis of Machine Learning Models in Financial Market Predictions.Expert Systems with Applications, 42(8), 3070-3081.
[12] Bouri, E., Molnár, P., Azzi, G., Roubaud, D., & Hagfors, L. I. (2017).On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?Finance Research Letters, 20, 192-198.
[13] Fischer, T., & Krauss, C. (2018).Deep learning with long short-term memory networks for financial market predictions.European Journal of Operational Research, 270(2), 654-669.
[14] Gandal, N., Hamrick, J. T., Moore, T., & Oberman, T. (2018).Price manipulation in the Bitcoin ecosystem.Journal of Monetary Economics, 95, 86-96.
[15] Corbet, S., Lucey, B., Urquhart, A., &Yarovaya, L. (2019).Cryptocurrencies as a financial asset: A systematic analysis.International Review of Financial Analysis, 62, 182-199.
[16] Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., & others. (2020).NumPy: Fundamental Package for Scientific Computing.Nature, 585(7825), 357–362. Retrieved from https://numpy.org/
[17] Shen, D., Urquhart, A., & Wang, P. (2020).Does Twitter predict Bitcoin?Economics Letters, 186, 108789.
[18] Jain, R. (2021).Building Machine Learning Applications with Streamlit. Packt Publishing.
[19] Binance API Documentation – Binance. (n.d.). Retrieved from https://binance-docs.github.io/apidocs/spot/en/.
[20] Streamlit: Open-Source App Framework for Machine Learning and Data Science – Streamlit Inc. (n.d.). Retrieved from https://docs.streamlit.io/.
[21] Pandas: Data Analysis Library. Retrieved from https://pandas.pydata.org/.



INDEX COPERNICUS

10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY