PRICING BEHAVIOR AND LIQUIDITY IN THE CRYPTOCURRENCY MARKET: AN EMPIRICAL ANALYSIS WITH REFERENCE TO BITCOIN AND ALTCOINS

A Thesis

Submitted for the award of the degree of

Doctor of Philosophy (Ph.D.)
in
Computational Finance

Submitted by

BHASKAR TRIPATHI

(Registration number: 951610003)

Under the Guidance of

Dr. Rakesh Kumar Sharma (Associate Professor)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

SCHOOL OF HUMANITIES AND SOCIAL SCIENCES

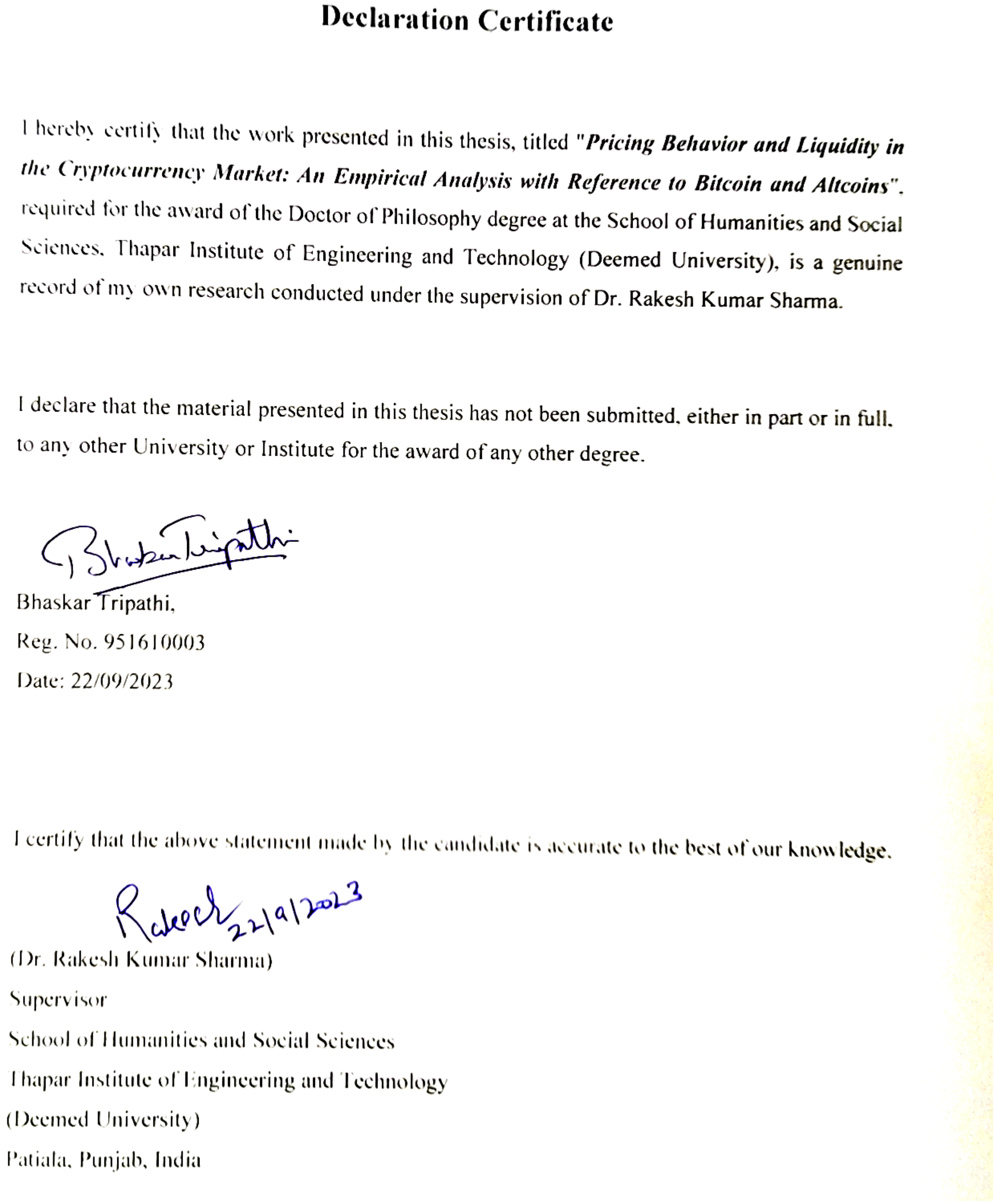
Thapar Institute of Engineering and Technology

(Deemed to be University)

PATIALA, PUNJAB (INDIA)

September 2023

Declaration Certificate



Declaration Certificate

I hereby certify that the work presented in this thesis, titled "***Pricing Behavior and Liquidity in the Cryptocurrency Market: An Empirical Analysis with Reference to Bitcoin and Altcoins***", required for the award of the Doctor of Philosophy degree at the School of Humanities and Social Sciences, Thapar Institute of Engineering and Technology (Deemed University), is a genuine record of my own research conducted under the supervision of Dr. Rakesh Kumar Sharma.

I declare that the material presented in this thesis has not been submitted, either in part or in full, to any other University or Institute for the award of any other degree.

Bhaskar Tripathi

Bhaskar Tripathi,
Reg. No. 951610003
Date: 22/09/2023

I certify that the above statement made by the candidate is accurate to the best of our knowledge.

Rakesh 22/9/2023

(Dr. Rakesh Kumar Sharma)
Supervisor
School of Humanities and Social Sciences
Thapar Institute of Engineering and Technology
(Deemed University)
Patiala, Punjab, India

This thesis is humbly dedicated to my beloved parents,
my spiritual master, Shree Ram Ji,
and to the countless researchers and authors whose wisdom and discoveries have shaped my understanding

Acknowledgements

I express my gratitude to the almighty, whose grace has enabled me to complete this work. I extend my heartfelt appreciation to those who have supported me throughout this challenging, yet rewarding, journey.

First and foremost, I express profound gratitude to my supervisor, Dr. R.K. Sharma, for his invaluable guidance and unwavering support throughout this journey. His expertise in research methodologies has critically influenced my academic objectives. His mentorship fostered an environment that encouraged exploration and creative thinking, providing clarity and direction in my pursuit of knowledge.

I would also like to extend my heartfelt thanks to Prof. Dr. Kaustuv Roy for the opportunity he offered for presenting my pre-submission work.

Dr. A.N. Shah's thought-provoking and insightful queries deserve a special mention. His ability to shift my perspective and encourage critical reevaluation of every aspect of my work has significantly contributed to my intellectual growth and understanding as a researcher.

Dr. Gurvinder Kaur for her encouragement and for illuminating new pathways of thought.

I express my gratitude to Dr. Ravi Kiran for providing me the opportunity to engage in the rigorous, yet enriching, coursework.

The invaluable support and scholarly guidance offered by the experienced doctoral committee members - Dr. Gurvinder Kaur, Dr. Rakesh Sharma, Dr. A.N. Sah, Dr. Harjot Singh (LMTSOM) have been instrumental in refining my research from various perspectives. Their valuable suggestions, thought-provoking comments, and challenging questions have not only broadened my research horizons but also strengthened my resolve.

I am sincerely thankful to Dr. Rudra Rameshwar for his pragmatic advice on conducting research and presenting findings. His thoughtful strategies have subtly refined my research approach, for which I express my heartfelt gratitude.

My acknowledgment would be incomplete without thanking the dedicated team at Nava Nalanda Central Library. Their steadfast assistance in providing access to several journals and the library's exceptional facilities have been instrumental in the progression of my research.

I extend heartfelt appreciation to my mother. Amidst the profound loss of my father, she emerged as a rock of support, both for my personal journey and my academic pursuits. In remembrance of my late father, a distinguished scientist with notable contributions to Microbiology, his impact has left an enduring imprint on my scholarly journey. The inspiration I garnered from his profound accomplishments has been instrumental in my pursuit of this Ph.D. His legacy as a meticulous critic continues to guide my academic exploration.

I would like to express my gratitude to my sister, whose firm belief in my capabilities has been a source of constant encouragement. Emulating the roles of both sister and mother, she has been an instrumental figure in my journey, reinforcing my resolve when it faltered and nurturing my academic growth.

My heartfelt appreciation goes to my wife, whose steadfast support has been invaluable throughout this journey. A special mention goes to my two young sons (Vitharya and Krithvik), whose innocent smiles have been a source of joy amidst my academic pursuits.

I would like to express my sincere gratitude to all the reviewers of this work. Your insights and feedback have been instrumental in refining this research. Lastly, for those not individually named here, yet who have offered their support behind the scenes, I deeply appreciate the value of your contributions.

Abstract

The thesis investigates the pricing dynamics and liquidity in the cryptocurrency market, focusing on Bitcoin and select Altcoins. It examines exchange price trends, forecasts future exchange prices, identifies the underlying factors of price formation, highlights discrepancies across various cryptocurrency exchanges, and compares liquidity of these exchanges with traditional stock exchanges.

Existing cryptocurrency market literature reveals several challenges arising from significant volatility and nonlinear price dynamics. Traditional econometric and statistical methods often encounter limitations during extreme market cycles prevalent in the cryptocurrency market, owing to their inherent linear assumptions in an essentially complex market structure. While past literature suggests that machine learning and deep learning models enhance forecasting accuracy, they suffer from a lack of transparency and explainability. In the first objective of this thesis, we address these challenges by analyzing the statistical properties of cryptocurrency price trends, assessing their temporal evolution, and examining the nature of data across various market periods. Before proceeding to data preprocessing, we utilize a high-dimensional multivariate dataset to broaden the range of factors potentially influencing the prices. Subsequently, given the intricacies observed within the data, we introduce a distinctive data preprocessing mechanism. We use a combination of signal-processing techniques to enhance data quality for effective modeling. A novel three-step feature selection method is proposed within this stage, ensuring efficient dimension reduction. The primary aim of this objective is to develop a highly accurate forecasting model that is robust to different market phases and varied short-term forecasting horizons. Accordingly, this research introduces a flexible and explainable financial forecasting architecture that ensures that the resulting predictions are interpretable and tailored for volatile asset market phases. Analysis using Change Point and Multifractal Detrended Fluctuation Analysis revealed that cryptocurrency prices follow non-linear trends with extreme price changes and asymmetric price distributions. Volatility showed significant fluctuations at multiple points, and the persistence of multifractality across daily and weekly scales underscored the complex nature of cryptocurrency price movements. These insights highlighted the need for an advanced forecasting model capable of capturing these dynamics. In response, we proposed a forecasting approach, combining signal processing with hybrid neural network models that utilized fundamental and technical indicators under different market conditions. This approach demonstrated superior performance compared to the existing state-of-the-art, conclusively demonstrating that accurate cryptocurrency price prediction requires sophisticated, multifaceted modeling techniques.

In the second objective of this research, we investigate the price formation factors of Bitcoin and Altcoins. While prior research primarily centers on Bitcoin, this study broadens its horizon by emphasizing major Altcoins, such as Ethereum, Ripple, Litecoin, and DASH. We also aim to understand how these Altcoins' price formation determinants differ from those of Bitcoin. Our analysis pivots on two fundamental research questions. Firstly, we aim to identify the core factors influencing Bitcoin and Altcoin prices. Secondly, we seek to ascertain how these influencing factors shift across distinct market phases and differ for each coin. Recognizing the challenge posed by varied influences on price formation, such as macroeconomic indicators, supply-demand dynamics, technological factors, and global financial trends, this study broadens

the scope beyond prior research that has often been limited to a subset of these factors. Moreover, acknowledging the standard critique of advanced predictive models regarding their lack of interpretability, our study incorporates the SHAP (SHapley Additive exPlanations) technique. SHAP Analysis of Bitcoin and Altcoins identified key factors influencing price formation, such as blockchain metrics, global financial indicators, technical analysis tools, and public sentiment. Bitcoin's price drivers evolved from blockchain-centric in early stages to market-related and operational efficiency in later stages. For Altcoins, including Ethereum, Ripple, Litecoin, and DASH, distinct factors like Bitcoin's price, social media sentiment, and blockchain activities were significant. This highlights the specific and shifting influences in cryptocurrency price dynamics across various market periods.

The third objective of this thesis focuses on identifying factors responsible for pricing inconsistencies across major cryptocurrency exchanges, namely Binance, Kraken, and Coinbase. Despite the 24/7 operations, global accessibility, and immediate access to pricing data of cryptocurrencies, actual market scenarios reveal notable pricing inconsistencies across different platforms. Despite the homogeneous nature of assets like Bitcoin and Ethereum, the market often deviates from the Law of One Price, an economic principle that suggests identical assets should have the same price across all markets. This phenomenon of price inconsistencies across different cryptocurrency exchanges, intrigues researchers and practitioners, and prompts further exploration. This research investigates the causes behind these inconsistencies, considering each exchange's unique operational characteristics, geographical locations, regulatory compliances, and market dynamics. In pursuit of this objective, the research seeks answers to two essential research questions: 1) What factors contribute to these price inconsistencies across exchanges? 2) How does the influence of these factors vary among different platforms? This analysis yields crucial findings that assist arbitrageurs, traders, and investors in making informed decisions about where to trade most profitably, ultimately enriching the knowledge base of cryptocurrency pricing dynamics. Our analysis identified key drivers of these discrepancies, including bid-ask spread, social media influence, especially Twitter sentiment, and various technical indicators like maker fees and Google Trends data. Each exchange displayed distinct influencing factors: Kraken's prices were significantly impacted by bid-ask spread and maker fees, Coinbase Pro's by Google Trends and Williams %R, and Binance's by transaction volume and market sentiment. This variation in influential factors across platforms illustrates the unique market dynamics and user behaviors characterizing each exchange, offering crucial insights for understanding pricing mechanisms in the cryptocurrency landscape.

The fourth objective of this thesis examines the liquidity of cryptocurrency exchanges compared to traditional stock markets. Given the integral role of liquidity in ensuring market efficiency and stability, it becomes essential to understand how cryptocurrency exchanges fare in relation to traditional stock markets. Few studies in past literature directly compare the liquidity dynamics of cryptocurrency platforms and traditional stock markets. We employ the Martin Liquidity Index (MLI) as the baseline measure for our comparative analysis. Alongside, we evaluate liquidity through four other established measures: Amihud's Illiquidity Ratio (AIR), AR Bid-Ask Spread, Roll's Covariance Liquidity Estimator, and the CS spread estimator. This multifaceted approach guides investment decisions and sets foundational liquidity benchmarks for the evolving cryptocurrency market. We specifically address two primary questions: 1) How does liquidity in traditional stock markets, as measured by the MLI, compare to major cryptocurrency exchanges? 2) Which of the leading cryptocurrency exchanges exhibits the highest liquidity, according to the MLI? The investigation aims to enhance understanding of liquidity dynamics, providing a framework to measure the maturation and effectiveness of cryptocurrency exchanges. This research aims to identify the most liquid cryptocurrency exchanges and assess how their liquidity compares to traditional stock exchanges, determining if they are more, equally, or less liquid. It demonstrates a consistent method for evaluating their market efficiency relative to traditional stock indices. The investigation, revealed that traditional stock markets, including NYSE, NASDAQ, NIFTY, and BSE SENSEX, generally have higher liquidity than major cryptocurrency exchanges. Among the examined cryptocurrency exchanges, Binance displayed the highest liquidity levels. This analysis confirms that traditional stock markets surpass cryptocurrency exchanges in liquidity, with Binance leading among the latter.

Overall, this research deepens our understanding of cryptocurrency markets by investigating pricing dynamics, forecasting, liquidity, and exchange discrepancies. Firstly, we present a flexible architecture designed for price forecasting using a high-dimensional, multivariate dataset. This architecture handles the non-linearity in cryptocurrency prices by employing advanced data pre-processing and signal-processing methods. Additionally, by integrating both technical and fundamental analysis, it efficiently predicts prices. Secondly, the study identifies key determinants that influence cryptocurrency prices. By understanding these factors, stakeholders can make more informed decisions, reducing the inherent risks associated with investments in this domain. Thirdly, we identify the factors for discrepancies in pricing across recognized cryptocurrency exchanges, Binance, Kraken, and Coinbase, and point to market inefficiencies, offering avenues for systematic arbitrage. Finally, by systematically comparing cryptocurrency exchanges' liquidity metrics with traditional stock markets, this research offers insights into both markets' relative liquidity and efficiency. This helps investors estimate the risk and return dynamics more proficiently. Together, these objectives enhance our understanding of cryptocurrency markets and provide tangible benefits to various stakeholders, from individual investors to financial institutions, practitioners, and academic researchers engaged in cryptocurrency market studies. The findings and outcomes of this work implications for stakeholders - investors, traders, researchers, and policymakers, aiding in strategic decision-making processes.

Table of Contents

Declaration Certificate **ii**
Acknowledgements **iv**
Abstract **vi**
List of Figures **xi-xii**
List of Tables **xiii-xiv**
Chapter 1. Introduction **1-13**
1.1 Cryptocurrency Price Trends and Market Evolution 3
1.2 Factors Influencing the Price of Cryptocurrencies 5
1.3 Variations in Cryptocurrency Pricing Across Different Exchanges 6
1.4 Liquidity Analysis of Cryptocurrency Markets and Traditional Stock Markets 7
1.5 Objectives of the Research 9
1.6 Research Questions and Hypothesis 9
1.7 Scope of the work 12
1.8 Thesis Structure and Outline 12

Chapter 2. Literature Review **14-39**
2.1 Cryptocurrency Price Trends and Market Evolution 14
2.1.1 Pre-Mount Gox Crash Period history of price swings 15
2.1.2 Price Trends Post Mount Gox Crash and Pre-2017 16
2.1.3 Price Trends between 2017 to 2021 16
2.1.4 Post 2021 Period history of price swings 17
2.1.5 Past Works 17
2.2 Price Formation Factors of Cryptocurrencies 26
2.3 Factors Responsible for Price Inconsistencies across Cryptocurrency Exchanges 31
2.4 Liquidity Analysis of Cryptocurrency Markets and Traditional Stock Indices 35

Chapter 3. Data and Research Methodologies **40-87**
3.1 Data 40
3.1.1 Dataset for Objectives 1 and 2 40
3.1.2 Dataset for Objective 3 44
3.1.3 Dataset for Objective 4 45
3.2 Research Methods 47
3.2.1 Objective 1 47
3.2.1.1 Analytical Methods for Investigating Cryptocurrency Price Patterns 47
3.2.1.2 The Proposed Cryptocurrency Price Forecasting Framework 53
3.2.1.3 Experimental Setup and Theoretical Background of the Proposed Framework 57
3.2.2 Objective 2 73
3.2.2.1 Explainable AI Framework to Determine the Price Formation Factors 73
3.2.2.2 DeepSHAP Based Feature Importance 74
3.2.3 Objective 3 76

3.2.3.1 Identifying Factors for Price Inconsistencies in Cryptocurrency Exchanges 76
3.2.4 Objective 4 79
3.2.4.1 Multifaceted Liquidity Assessment 79
3.2.4.2 Proposed Framework for Liquidity Analysis 80
3.2.4.2.1 Experimental Settings 80
3.2.4.2.2 Liquidity Calculation Methods 81

Chapter 4. Results and Discussion 82-146
4.1 Objective 1: Trends and Forecasts of Bitcoin and Altcoin Prices 88
4.1.1 Results for Objective 1 88
4.1.1.2 Forecasting Results 109
4.1.1.2.2 Comparing Our Results with Previous Literature 111
4.2 Objective 2: Price Formation Factors of Bitcoin and Altcoins 111
4.2.1 Results for Objective 2 111
4.2.2 Main Contributions and Novelty 122
4.3 Results for Objective 3: Factors for Price Inconsistencies Across Exchanges 123
4.3.1 Descriptive Statistics of Prices and Price Inconsistencies 123
4.3.2 Absolute and Relative Price Discrepancies of Each Exchange 127
4.3.3 Model Performance Metrics and SHAP Analysis Results 130
4.3.4 Implications for Stakeholders: Traders, and Arbitrageurs 138
4.4 Objective 4: Liquidity Comparison 139
4.4.1 Liquidity Analysis: Metrics and Comparisons Across Exchanges 139
4.4.2 Implications for Traders, Researchers, and Arbitrageurs. 146

Chapter 5. Conclusions and Future Work 147-162
5.1 Conclusion and Summary of Main Findings: Objective 1 147
5.1.1 Recap of Objective 1 148
5.1.2 Major Findings and Research Outcomes 148
5.1 Conclusion and Summary of Main Findings: Objective 2 153
5.2.1 Recap of Objective 2 153
5.2.2 Key Factors Influencing Price Formation of Bitcoin and Altcoins 153
5.2.3 Research Question 2.2 156
5.3 Conclusion and Summary of Main Findings: Objective 3 157
5.3.1 Recap of Objective 3 157
5.4 Conclusion and Summary of Main Findings: Objective 4 158
5.4.1 Recap of Objective 4 159
5.4.2 Research Questions and Major Findings 159
5.5 Implications of this Research 160
5.6 Limitations of this Research and Future Research Directions 161
References 163-172
Appendix A 173-176
Appendix B 177
B1.1 List of Publications and Patents 178

List of Figures

		Page No.
1	Figure 1. Historical Closing Prices of Bitcoin and Ethereum, the Two Leading Coins in the Cryptocurrency Space (Normalized for Comparison)	15
2	Figure 2. Systematic Workflow of the Proposed Cryptocurrency Price Forecasting Framework	54
3	Figure 3 illustrates the schematic structure of an ANN, which includes an input layer, three hidden layers, and an output layer.	65
4	Figure 4. Schematic diagram of a BiLSTM Network	67
5	Figure 5. Schematic Architecture of a CNN-BiLSTM Network with Self Attention	68
6	Figure 6. Schematic representation of the analytical framework for determining the price formation factors of Bitcoin and Altcoins	74
7	Figure 7. Workflow for Analyzing Cryptocurrency Price Inconsistencies.	76
8	Figure 8. Workflow for calculating Martin Liquidity Index (MLI) of different cryptocurrency exchanges	81
9	Figure 9 (a). Change Point analysis of the exchange prices of Bitcoin.	92
10	Figure 9 (b). Analytical depiction of changepoints in the Ethereum time series.	93
11	Figure 9 (c). Analytical depiction of Change Points in the DASH time series.	94
12	Figure 9 (d). Analytical depiction of Change Points in the Litecoin (LTC) time series.	94
13	Figure 9 (e). Analytical depiction of Change Points in the Ethereum time series.	95
14	Figure 9 (f). Multifractal Detrended Fluctuation Analysis (MFDFA) of Bitcoin and Altcoins.	97
15	Figure 10. Comparison of predicted and actual closing prices of Bitcoin for Interval 4 using the best-performing model (testing dataset).	103
16	Figure 11. Comparison of predicted and actual closing prices of Ethereum for Interval 4 using the best-performing model (testing dataset).	103
17	Figure 12. Comparison of predicted and actual closing prices of Ripple for Interval 4 using the best-performing model (testing dataset).	104
18	Figure 13. Comparison of predicted and actual closing prices of Litecoin for Interval 4 using the best-performing model (testing dataset).	104
19	Figure 14. Comparison of predicted and actual closing prices of DASH for Interval 4 using the best-performing model (testing dataset).	105
20	Figure 15. Model loss progression for Bitcoin over 3-day, 5-day, and 7-day forecast horizons (testing dataset).	107
21	Figure 16. Model loss progression for Ethereum over 3-day, 5-day, and 7-day forecast horizons (testing dataset).	107
22	Figure 17. Model loss progression for Ripple over 3-day, 5-day, and 7-day forecast horizons (testing dataset).	108
23	Figure 18. Model loss progression for Litecoin over 3-day, 5-day, and 7-day forecast horizons (testing dataset).	108

24 Figure 19. Model loss progression for DASH over 3-day, 5-day, and 7-day forecast horizons (testing dataset). 109
25 Figure 20. Summary plot and Bar plot to illustrate the main determinants of Bitcoin’s price formation. 112
26 Figure 21. Summary plot and Bar plot obtained from SHAP analysis illustrate the Global Feature Importance of Bitcoin’s Interval 2 (29-Apr-13 to 30-Apr-17). 113
27 Figure 22. Global Feature Importance for Bitcoin’s Interval 3 (29-Apr-13 to 31-Dec-17), showing the shift in feature influence compared to previous intervals. 113
28 Figure 23. Global Feature Importance summary plot derived from SHAP values for the best performing CNN-BiLSTM model with Fundamental and Technical Indicators as the input dataset for Bitcoin’s Interval 4 (2-Jan-17 to 2-Apr-23). 114
29 Figure 24. Ethereum’s Interval 4 (2-Jan-17 to 2-Apr-23) Feature Importance. 114
30 Figure 25. Ripple’s Interval 4 (2-Jan-17 to 2-Apr-23) Feature Analysis. 115
31 Figure 26. Litecoin’s Interval 4 (2-Jan-17 to 2-Apr-23) Feature Breakdown. 115
32 Figure 27. DASH’s Interval 4 (2-Jan-17 to 2-Apr-23) Feature Insights. 116
33 Figure 28. SHAP Feature Importance Plots for Price Inconsistencies Across Major Cryptocurrency Exchanges: (a) Kraken, (b) Coinbase, and (c) Binance 135
34 Figure 29 (a). Global Feature Importance for Kraken 134
35 Figure 29 (b). Global Feature Importance for Coinbase Pro 135
36 Figure 29 (c). Global Feature Importance for Binance (Global). 135
37 Figure 30 (a) and 30 (b). Radar chart representation of the Martin Liquidity Index (MLI) and Amihud Illiquidity Measure values across various exchanges for all three intervals. 142
38 Figure 30 (c) and 30 (d). Radar chart representation of the Average Relative Bid-Ask Spread and CS Estimator values across different exchanges for three distinct intervals. 142
39 Figure 30 (e). Radar chart representation of Roll's Estimator for the exchanges, presented with a scaled axis to accommodate the range of values, marked at 1×10−9. 143

List of Tables

1. Table 1. A chronological review of past works on cryptocurrency price forecasting. 19

2. Table 2.1 Summary of the Fundamental and Technical indicators used as independent variables for Objective 1 and Objective 2. 41

3. Table 2.2 A descriptive overview of Blockchain variables for Bitcoin and Altcoins (Ethereum, Ripple, Litecoin and DASH) used as independent variables for Objective 1 and Objective 2.

4. Table 3. Data intervals considered for Bitcoin (BTC) 44

5. Table 4. Data intervals considered for Ethereum (ETH) 44

6. Table 5. Data intervals considered for Litecoin (LTC) 44

7. Table 6. Data intervals considered for Ripple (XRP) 44

8. Table 7. Data intervals considered for Dash 44

9. Table 8. Independent variables utilized for evaluating discrepancies in cryptocurrency prices across major exchanges 45

10. Table 9. Data extraction intervals for liquidity comparison across exchanges. 46

11. Table 10. Hyperparameter search space for Bayesian Optimization in deep learning model tuning 70

12. Table 11. Descriptive statistics for Change Points obtained from PELT algorithm for Bitcoin (BTC) 88

13. Table 12. Descriptive statistics for Change Points obtained from PELT algorithm for Ethereum (ETH) 89

14. Table 13. Descriptive statistics for Change Points obtained from PELT algorithm for DASH 89

15. Table 14. Descriptive statistics for Change Points obtained from PELT algorithm for Litecoin (LTC) 90

16. Table 15. Descriptive statistics for Change Points obtained from PELT algorithm for Ripple (XRP) 90

17. Table 16. Mean results of 20 runs for next day closing price prediction of Bitcoin (BTC) for all four intervals 99

18. Table 17. Mean results of 20 runs for next day closing price prediction of Ethereum (ETH) for all four intervals 100

19. Table 18. Mean results of 20 runs for next day closing price prediction of Ripple (XRP) for all four intervals 100

20. Table 19. Mean results of 20 runs for next day closing price prediction of Litecoin (LTC) for all four intervals 100

21. Table 20. Mean results of 20 runs for next day closing price prediction of DASH coin for all four intervals 100

22. Table 21. The mean forecast results for Bitcoin's closing price over 'N' days across 20 runs. 105

23. Table 22. The mean forecast results for Ethereum's closing price over 'N' days across 20 runs. 106

24. Table 23. The mean forecast results for Ripple’s closing price over 'N' days across 20 runs. 106
25. Table 24. The mean forecast results for Litecoin's closing price over 'N' days across 20 runs. 106
26. Table 25. The mean forecast results for DASH's closing price over 'N' days across 20 runs. 106
27. Table 26. Descriptive statistics for closing prices of Bitcoin (BTC) 125
28. Table 27. Descriptive statistics for closing prices of Ethereum (ETH) 125
29. Table 28. Descriptive statistics for closing prices of Ripple (XRP) 125
30. Table 29. Descriptive statistics for closing prices of Litecoin (LTC) 125
31. Table 30. Descriptive statistics for closing prices of DASH 126

32. Table 31. Descriptive Statistics of Absolute and Relative Price Discrepancies (referred to as "Spread" in this and subsequent tables) across exchanges for Bitcoin (BTC). 128
33. Table 32. Descriptive Statistics of Absolute and Relative Price Discrepancies Across Exchanges for Ethereum (ETH). 128
34. Table 33. Descriptive Statistics of Absolute and Relative Price Discrepancies Across Exchanges for Ripple (XRP). 128
35. Table 34. Descriptive Statistics of Absolute and Relative Price Discrepancies Across Exchanges for Litecoin (LTC). 128
36. Table 35. Descriptive Statistics of Absolute and Relative Price Discrepancies Across Exchanges for DASH. 128
37. Table 36. List of Selected Features from the Three-Step Feature Selection Process. 131
38. Table 37. Hyperparameters suggested by Bayesian Optimization 132
39. Table 38. Price Inconsistency Prediction Errors Across Major Cryptocurrency Exchanges: Kraken, Coinbase, and Binance. 132
40. Table 39: SHAP Feature Importance for Price Inconsistencies Across Major Cryptocurrency Exchanges: Kraken, Coinbase, and Binance. 133
41. Table 40. Liquidity Measures Across Major Cryptocurrency and Traditional Markets for Interval 1. 141
42. Table 41. Liquidity Measures Across Major Cryptocurrency and Traditional Markets for Interval 2 141
43. Table 42. Liquidity Measures Across Major Cryptocurrency and Traditional Markets for Interval 3 141
44. Table 43. Evaluation of Hypothesis for Research Question 1.1 150
45. Table 44. Evaluation of Hypothesis for Research Question 1.2 152
46. Table 45. Evaluation of Hypothesis for Research Question 1.3 153
47. Table 46. Evaluation of Hypothesis for Research Question 5.4.1 161

Chapter 1

Introduction

Cryptocurrencies are digital assets intended to act as a medium of exchange within a decentralized network structure (Baur et al., 2018). Unlike traditional fiat currencies, they operate on decentralized platforms powered by Blockchain technology. This technology serves as the underlying infrastructure for cryptocurrencies and a variety of other applications, offering benefits such as enhanced security, privacy, and scalability on an immutable, distributed ledger system (Misra et al., 2023).

Each transaction is bundled with other transactions to create a 'block.' These blocks are cryptographically secured and linked in a linear, chronological order, forming a 'chain'; hence, the name 'Blockchain.' The Blockchain structure ensures the immutability and transparency of all transactions (Ammous, 2015). Every transaction within a block is authenticated by network participants, commonly known as 'miners,' who utilize cryptographic algorithms for this purpose. These miners confirm the legitimacy of transactions, effectively preventing double-spending - spending the same cryptocurrency units more than once (Y. Chen, 2018). Once transactions are successfully authenticated, the new block — with all its validated transactions — is appended to the existing chain of blocks. This 'proof-of-work' process safeguards the Blockchain from potential tampering and unauthorized modifications, rendering the creation of additional units or alteration of past transactions computationally unfeasible. Enabled by this technology, the transactions can take place directly between involved parties over "peer-to-peer" networks, effectively eliminating the need for traditional intermediaries such as banks or payment gateways (Nakamoto, 2008).

Bitcoin, the world’s first cryptocurrency, was introduced by a pseudonymous entity, Satoshi Nakamoto, in 2008 (Nakamoto, 2008). The primary innovation Bitcoin brought forth was a solution to the "double-spending" problem, an issue peculiar to digital currencies. Double-spending pertains to the risk of a single digital token being spent more than once, made possible due to the replicable nature of digital information. In a blockchain, every transaction is time-stamped and linked to the previous transaction, ensuring that every coin spent has a traceable history, which prevents double-spending. To address the double-spending problem, cryptocurrencies, such as Bitcoin, adopt a system known as "proof-of-work" (PoW) (Patel et al., 2022). This system is a consensus algorithm integral to the security and validity of the decentralized nature of cryptocurrencies. In PoW, participants, known as miners, engage in complex computations, requiring significant processing power and energy, to solve intricate mathematical puzzles. The PoW mechanism operates under the premise that the computational work required is costly and time-consuming. As a result, it is generally unfeasible for an entity to perform the work necessary to validate fraudulent transactions or alter past transactions. The first miner to solve the puzzle gains the right to add a new block of valid transactions to the blockchain network. This achievement is accompanied by a reward in the form of cryptocurrency, which serves as an incentive for miners to engage in the strenuous PoW process (Nakamoto, 2008).

In contemporary finance, cryptocurrencies serve primarily as speculative financial assets, with Bitcoin, Ethereum, and their counterparts attracting significant attention due to their high return potential (A. D. Lee et al., 2020). Despite inherent volatility and fluctuating values, these digital assets present fertile ground for trading activity, driven primarily by market speculation. Investors closely monitor market trends and news surrounding cryptocurrencies, as the smallest shift can trigger dramatic price fluctuations. The decentralized nature of these digital currencies, coupled with their limited supply, contribute to their allure. The emergence of numerous Altcoins has further diversified the cryptocurrency market, expanding the range of investment opportunities (Ong et al., 2015). As technology advances, the Blockchain technology underlying cryptocurrencies is gaining recognition for its potential applications beyond finance, such as supply chain management, maintaining land records and voting systems (Ali et al., 2020). Cryptocurrencies, while predominantly used as financial assets, also provide multifaceted utility. Their capacity to bypass conventional banking and financial inefficiencies enables rapid, low-cost transactions. In economically unstable regions, cryptocurrencies may serve as alternative mechanisms for wealth preservation, resistant to external interference and censorship. An increasing number of sectors are accepting cryptocurrencies for goods and services, extending their functionality beyond pure speculation (Härdle et al., 2020).

With the emergence of 'Altcoins,' such as Ethereum, Ripple, Litecoin, and Dash, the cryptocurrency market has experienced notable growth and diversification since its inception. Each Altcoin possesses distinctive features and applications, expanding the investor's inventory and diversifying the market. The maturation of the cryptocurrency landscape has ushered in the development of stablecoins such as Tether (USDT), USD Coin (USDC), Binance USD (BUSD), Dai (DAI), and TrueUSD (TUSD). These digital financial instruments present a strategic solution to the extensively reported volatility pervasive in the cryptocurrency market. With their values linked to reserves of traditional currencies, most frequently the US dollar, these stablecoins provide an interface between traditional fiat currencies and digital assets. Consequently, they provide a more reliable investment path within the multifaceted cryptocurrency space (KAYAHAN et al., 2022).

The rise of decentralized finance (DeFi) has expanded the investment possibilities within the cryptocurrency domain. DeFi platforms offer an array of financial services independent of traditional banking intermediaries, providing increased opportunities for historically marginalized individuals (Schueffel, 2021). Concurrently, non-fungible tokens (NFTs) extend the scope of cryptocurrency investments, creating new opportunities for investment, monetization, and wealth creation. Additionally, the increasing number of non-fungible tokens (NFTs) expands the scope of cryptocurrency investments. By facilitating the tokenization of different digital assets, such as artwork, collectibles, and virtual real estate, the NFT market creates new opportunities for investment, monetization, and wealth creation (Ali et al., 2023). As cryptocurrency adoption increases, they are increasingly incorporated into traditional financial services. Governments are investigating central bank digital currencies (CBDCs) as large corporations begin to accept them as payment (Agur et al., 2022). As cryptocurrencies continue to evolve, they are bound to elicit considerable interest from various quarters, including investors, governments, and institutions, emphasizing their significance in the global financial landscape.

1.1 Cryptocurrency Price Trends and Market Evolution

The historical prices of cryptocurrencies have been characterized by extreme volatility over the past decade. From Bitcoin's inception in 2009, its price has seen striking booms and significant crashes, with similar trends also seen in other major Altcoins. Starting in 2017, Bitcoin's heightened recognition as a premier asset significantly influenced the cryptocurrency market, sparking amplified interest among investors. This wave of acknowledgement not only escalated Bitcoin's demand but also propelled the prominence and value of other cryptocurrencies such as Ethereum, Ripple, and Litecoin. Consequently, the cryptocurrency market capitalization reached its peak in November, 2021 (Younis Masiha, 2022). However, the historical volatility of Bitcoin is considerable, marked by substantial price swings, including notable declines within a single day after achieving record highs (Carbó & Gorjon, 2022a). This volatility, while being a risk factor, has also attracted speculators and traders who seek to profit from the price swings. Despite the volatile history, the overall trend for these cryptocurrencies has been upward, with an increasing number of investors and adoption rates over the past decade.

The extreme volatility in the cryptocurrency market significantly impacts investment strategies and market stability (Pichl & Kaizoji, 2017). On one hand, it creates opportunities for high returns for risk-tolerant investors and traders. On the other hand, it poses risks of significant losses. It also makes the task of portfolio optimization more complex as it becomes challenging to forecast the prices accurately (Schellinger, 2020). From the perspective of market stability, the extreme price swings can lead to periods of market stress, impact liquidity conditions, and even cause systemic risks in extreme cases (Al-Yahyaee et al., 2020). This complexity poses unique challenges for investors and traders in the cryptocurrency market. Diversification across different cryptocurrencies, using derivative instruments like futures and options, and employing automated trading algorithms are some common strategies. Cryptocurrency traders also use technical analysis techniques to understand price trends and identify trading opportunities (Ntakaris et al., 2020). Risk management strategies, including setting stop-loss limits and position sizing, are crucial to protect against extreme price swings.

Several factors contribute to price volatility in cryptocurrencies (Park & Chai, 2020). Market sentiments driven by news events, regulatory announcements, and macroeconomic developments can cause abrupt price changes (Huang et al., 2019). Additionally, the liquidity conditions in the cryptocurrency market can impact volatility. During periods of low liquidity, even small trades can cause significant price movements. The decentralized and global nature of cryptocurrencies also means that they trade round the clock, which can lead to abrupt price changes at any time. Moreover, the dynamic nature of the cryptocurrency market, characterized by evolving market cycles, changes in the public interest, advancements in Blockchain technology, and varying degrees of speculation and adoption, further complicates price prediction. The determinants influencing pricing in one phase of the market's evolution may not remain relevant in the subsequent phase, leading to investor uncertainty (Tripathi & Sharma, 2022). The inherent volatility of cryptocurrencies is fundamentally linked to their unique characteristics, including non-linearity, multi-modality, and a multi-fractal nature. Non-linearity refers to the unpredictable and disproportionate response of prices to market events. Multi-modality means that cryptocurrency price distributions exhibit multiple peaks, reflecting the possibility of sudden and extreme price changes. The multi-fractal (Lahmiri & Bekiros, 2018) and multi-scale (Wątorek et al., 2021) characteristics imply that price patterns repeat themselves on different time scales, from short-term fluctuations to long-term trends.

Given these complexities, the cryptocurrency market's dynamic and evolving nature unfolds opportunities and challenges for investors and businesses. One critical aspect of addressing these challenges is the ability to forecast prices accurately. The accuracy of such forecasts directly influences decision-making processes and investment strategies, emphasizing the need for a comprehensive and rigorous analysis of price trends. This analysis not only involves exploring the key statistical properties that shape the price patterns of Bitcoin and Altcoins but also analyzing how these properties change through different market phases. Incorporating technical and fundamental indicators in the modeling process is essential, as they collectively provide a comprehensive representation of market prices. These determinants encompass a spectrum of factors, from blockchain-related factors and macroeconomic data to market sentiment and technological advancements. Including these indicators in the input data, improves the likelihood that the model accurately mirrors market prices, leading to improved forecasting accuracy. Therefore, to carefully analyze the price trends for informed decision-making, we ask the following research questions:

Q1. What are the key statistical properties characterizing the price trends of Bitcoin and Altcoins?

Q2. How do these statistical properties change across different phases of the cryptocurrency market's evolution?

Q3. How do fundamental indicators, technical indicators, and their combination influence the predictive performance in forecasting the closing prices of Bitcoin, Ethereum, and other Altcoins?

1.2 Factors Influencing the Price of Cryptocurrencies

Bitcoin operates relatively distinctly from traditional economic structures, influenced partly by investors' speculative behavior. Nevertheless, it is not entirely disconnected from the real economy. Consequently, standard economic theories often struggle to capture all the factors influencing Bitcoin's price dynamics (Zhu et al., 2017). The price of Bitcoin is determined by the equilibrium between its supply and demand in the market. Supply governs the number of Bitcoins in circulation, dictating their scarcity or abundance in the market (Mayo & Elgazzar, 2022; Morillon & Chacon, 2022). On the other hand, demand is shaped by transactional demand as a medium of exchange and its employment in speculative trading (Blau, 2018; A. D. Lee et al., 2020).

Global financial development, captured by stock exchange indices, exchange rates, and oil prices, also plays a role in determining Bitcoin price. The cost of Bitcoin production, which resembles a competitive market, may represent a theoretical value around which market prices gravitate. Moreover, sentiment analysis performed on Twitter feeds, Google trends, and cryptocurrency news, which indicate the degree of public interest in cryptocurrencies, have positively affected their prices (Kayal & Rohilla, 2021). The mining cost anchors the market price, but as mining technology becomes efficient, the effect of mining difficulty on the market price diminishes over time. Given the dynamic nature of the cryptocurrency market, it is crucial to continually research and update the understanding of these price formation factors. The evolving market conditions of cryptocurrencies (Wątorek et al., 2023), their technological advancements, inherent multi-scale characteristics (Wątorek et al., 2021) and changing investor behaviors necessitate a periodic reassessment of these determinants to maintain accurate and relevant insights into cryptocurrency price formation.

The price formation of cryptocurrencies is a complex process influenced by various factors, including macroeconomic indicators, supply and demand dynamics, technological factors, and global financial trends. However, the unique characteristics of cryptocurrencies, such as their decentralized nature and the speculative behavior of investors, challenge the applicability of standard economic theories to explain their price formation fully. Furthermore, the rapidly evolving nature of the cryptocurrency market necessitates continuous research to keep pace with changing market conditions, technological advancements, and investor behaviors (W. Chen et al., 2021; Georgoula et al., 2015; Li & Wang, 2017). While existing research has explored individual or a small subset of factors influencing cryptocurrency prices, there is a need for a broader investigation that collectively examines these determinants (Patel et al., 2022). This necessitates a comprehensive study that concurrently investigates these diverse factors, providing a more nuanced understanding of cryptocurrency price formation. Despite the advancements in machine learning and deep learning models for price prediction, a significant research gap exists in the domain of model interpretability (Lundberg & Lee, 2017). Most of these works often lack transparency and explainability, making it difficult to understand the underlying factors driving the predictions. This is particularly relevant in the context of cryptocurrencies, given the large number of variables influencing price formation and the inherent complexity of the market. Incorporating explainable AI techniques (Islam et al., 2022; Shi et al., 2021) like SHAP enhances the interpretability of machine learning models by identifying the most important features for the model's predictions, thereby bridging the existing research gap in this domain. The distinct properties of cryptocurrencies and the dynamic nature of the cryptocurrency market make it challenging to apply traditional economic theories to explain their price formation (Saad et al., 2020). A detailed analysis that considers numerous factors influencing cryptocurrency prices is crucial for a deeper understanding of price formation factors. Additionally, there is a need for further research in model interpretability, which could be addressed by using explainable AI techniques to identify the key features driving the predictions.

Due to the complex nature of Bitcoin's decentralized peer-to-peer network and the influence of various internal and external factors on its price, there is a need for ongoing research to comprehensively understand the formation of Bitcoin's price. The research questions to understand this price formation factors of Bitcoin and various Altcoins are stated as follows:

Q1: What are the key factors that influence the price formation of Bitcoin and Altcoins?

Q2: How does price formation factors' impact on Bitcoin and Altcoins change across different market periods?

1.3 Variations in Cryptocurrency Pricing Across Different Exchanges

The dynamic phenomenon of price discovery in the cryptocurrency market, including Bitcoin and other digital assets, is significantly influenced by multiple factors. Among these factors are trading volumes, inflation rates, exchange rates, and regulatory changes, all of which contribute to the varying price points on different Bitcoin exchanges. The variability in information distribution among exchanges further shapes the price determination process, underlining the importance for traders and investors to stay updated on these dynamics (Akba et al., 2021). These include not just trading volumes and liquidity, but also regulatory changes and the emergence or closure of exchanges. Given the swiftly changing Bitcoin exchange ecosystem, continuous research and analysis are warranted. This research seeks to delve into the compelling area of price discrepancies across cryptocurrency exchanges, aiming to identify the contributing factors for better comprehension of price formation.

In the unique trading environment of the cryptocurrency market, characterized by 24/7 operation, global accessibility, and immediate availability of pricing data, one would anticipate the Law of One Price (LOOP) to hold. LOOP posits that identical assets, such as Bitcoin, Ethereum, and other cryptocurrencies, should have the same price across all markets, considering trading costs (Andrade et al., 2021). However, inconsistencies in Bitcoin prices across different exchanges contradict this expectation (Pieters & Vivanco, 2017), despite the identical nature of cryptocurrencies traded on various platforms.

Several factors specific to different exchanges, such as unique operational characteristics, geographical locations, and regulatory compliances, can significantly influence cryptocurrency prices. These factors lead to observed price inconsistencies across exchanges (Brauneis et al., 2022), an issue with substantial profitability implications for traders and investors (Koutmos & Wei, 2023). Despite its importance, research identifying these influencing factors remains limited, hindering a comprehensive understanding of price formation. Addressing this research gap, this investigation aims to identify and analyze the key elements contributing to price inconsistencies across major cryptocurrency exchanges. To achieve this, the research poses the following questions:
Q1: What factors contribute to price inconsistencies across major cryptocurrency exchanges (Binance, Kraken, Coinbase)?
Q2: How does the influence of these factors vary across different cryptocurrency exchanges (Binance, Kraken, Coinbase)?

1.4 Liquidity Analysis of Cryptocurrency Markets and Traditional Stock Markets

The liquidity of cryptocurrency exchanges, especially those with high trading volumes, is crucial to the market's continued stability and efficiency (Loi, 2017). Without sufficient liquidity, the market becomes susceptible to price manipulation and extreme volatility. This essential liquidity fosters a market environment that allows for seamless transactions and precise price discovery, hence attracting institutional investors. Such investors, with their large order sizes, enhance the stability and credibility of the market.

In addition, a liquid market reduces risks associated with holding illiquid assets during market stress, by enabling traders to rapidly enter and exit positions. It also narrows bid-ask spreads and reduces user transaction fees by promoting competition among exchanges (Lybek & Sarr, 2002). Another significant aspect of high liquidity is that it prevents substantial price shifts during transactions. Moreover, liquidity risk, a dynamic evaluation, can provide early indicators of market instability and expose processes leading to market crashes.

In cryptocurrency market, the liquidity of various cryptocurrency exchanges is varied suggesting a distinct trading environment in each platform (Brauneis et al., 2021a). Traditional stock markets typically demonstrate a greater degree of liquidity in comparison to cryptocurrency exchanges. This disparity in liquidity emphasizes the distinct characteristics inherent to each market. For example, the traditional stock markets are strictly regulated, mature, and highly liquid. In contrast, cryptocurrency exchanges are relatively nascent, unregulated, and subject to high volatility, often exhibiting lower liquidity levels (Cortez et al., 2021). As the cryptocurrency markets become more tightly integrated with the financial markets, it is necessary to investigate their market liquidity compared to conventional financial markets. Without enough practical understanding, traders may be obligated to use platforms that do not consistently offer optimal liquidity or pricing conditions. When the liquidity of a trading platform is inconsistent, it can create difficulties for traders aiming to execute trades at specific prices, especially when handling larger volumes. In such a scenario, the issue of slippage can become intensified. Slippage refers to the trading situation when the expected price of a trade differs from the actual price at which it is executed, which can subsequently affect the trade's profitability. This discrepancy can impact the profitability of the trade. In the evolving cryptocurrency landscape, traders face unique challenges due to the operational and functional differences between cryptocurrency and traditional exchanges (Brauneis et al., 2022). This makes a practical impact on their trading and profitability strategies. Operational and functional differences between cryptocurrency and traditional exchanges add another layer of complexity to trading, influencing the choice of trading platform with profound implications on cost, speed, and success of operations. In illiquid markets, large trades can significantly impact market price, leading to unfavorable trade execution. Further, transaction costs may increase due to wider bid-ask spreads, diminishing potential profits. Low liquidity markets can also result in delayed order fulfillment, disrupting traders dependent on swift execution to capitalize on short-term market movements. Moreover, low liquidity can severely limit the effectiveness of trading strategies that rely on quick and cost-effective trade of large volumes, such as high-frequency trading.

This understanding is vital as despite the growing prominence of cryptocurrencies, there is a significant gap in comprehending the factors that influence their liquidity. A comparison facilitated by a standardized measures such as Martin Liquidity Index (MLI), bid-ask spread, and Amihud’s illiquidity ratio among others can reveal operational differences between cryptocurrency and traditional stock exchanges, thereby informing investment decision-making. Such an analysis can also contribute to the development of liquidity benchmarks for the relatively new and rapidly evolving cryptocurrency market. Understanding the liquidity of leading cryptocurrency markets and traditional stock exchanges necessitates a thorough examination of various factors.

This research aims to delve into this aspect using the Martin Liquidity Index (MLI), where a higher value denotes lower liquidity. The MLI comparison provides a comprehensive understanding of liquidity patterns across diverse markets and acts as a benchmark for assessing the maturity and efficiency of cryptocurrency exchanges. Particularly, we need to discern how the liquidity of traditional stock markets (NYSE, NASDAQ, BSE Sensex) compares to that of major cryptocurrency exchanges (Binance, Kraken, Coinbase) as measured by the Martin Liquidity Index (MLI). Further, an exploration of which among the major cryptocurrency exchanges (Binance, Kraken, Coinbase) exhibits the highest liquidity as indicated by the Martin Liquidity Index (MLI) is crucial. Undertaking this analysis will offer valuable insights to a diverse array of stakeholders such as researchers, cryptocurrency traders, financial regulators, investors, miners, analysts, and portfolio managers. The study thereby intends to bridge existing gaps in research and contribute to a more sophisticated understanding of cryptocurrency market dynamics. To acquire an in-depth understanding of the complexities of cryptocurrency and traditional stock market liquidity, we pose the following research questions.

Q1: How does the liquidity of traditional stock exchanges (NYSE, NASDAQ, BSE Sensex) compare to that of major cryptocurrency exchanges (Binance, Kraken, Coinbase) as measured by the Martin Liquidity Index (MLI)?

Q2: Which of the major cryptocurrency exchanges (Binance, Kraken, Coinbase) exhibits the highest liquidity as measured by the Martin Liquidity Index (MLI)?

1.5 Objectives of the Research

The objectives of this research are stated as follows:

- **O1:** To analyze the trends of exchange prices of Bitcoin and Altcoins and to forecast future exchange prices.
- **O2:** To determine the factors that influence the price formation of Bitcoin and Altcoins.
- **O3:** To identify the factors responsible for price inconsistencies across different major cryptocurrency exchanges.
- **O4:** To compare the liquidity of cryptocurrency exchanges with that of different sizes of selected stocks using the Martin Liquidity Index (MLI).

Altcoins considered for the scope of this study are Ethereum (ETH), Ripple, Dash and Litecoin.

1.6 Research Questions and Hypothesis

Objective 1: *To analyze the trends of exchange prices of Bitcoin and Altcoins and to forecast future exchange prices*

- **Research Question 1.1:** What are the key statistical properties characterizing the price trends of Bitcoin and Altcoins?
- **H 1.1a:**
    - H1: The price trends of Bitcoin and Altcoins are non-linear, with distinct periods of growth and decline characterized by shifts in mean prices.
- **H1.1b:**
    - H1: The price trends of Bitcoin and Altcoins exhibit leptokurtosis, indicating a higher probability of extreme price changes during certain periods.
- **H1.1c:**
    - **H1**: The price trends of Bitcoin and Altcoins exhibit skewness, indicating asymmetry in price changes with periods of higher probabilities for both large price increases and decreases.

- **Research Question 1.2:** How do these statistical properties change across different phases of the cryptocurrency market's evolution?
    - **H1.2a:**
        - **H1**: The volatility of Bitcoin and Altcoins fluctuates significantly across different segments, reflecting the inherent risk and uncertainty in the cryptocurrency market.
    - **H 1.2b:**
        - **H1**: Cryptocurrencies exhibit complex scaling behavior, with multifractal properties in their price time series.

- **Research Question 1.3:** How do fundamental indicators, technical indicators, and their combination influence the predictive performance in forecasting the closing prices of Bitcoin, Ethereum, and other Altcoins?
    - **Hypothesis 1.3:**
    - **H1**: The models, when provided with a combination of Fundamental Indicators (FI) and Technical Indicators (TI) as input, yield superior forecasting accuracy for cryptocurrency closing prices compared to using either FI or TI alone with these models.

Objective 2: *To determine the factors that influence the price formation of Bitcoin and Altcoins.*

- **Research Question 2.1:** What are the key factors that influence the price formation of Bitcoin and Altcoins?
- **Research Question 2.2:** How does the impact of price formation factors on Bitcoin and Altcoins change across different market periods?

Given the complexity of the cryptocurrency market, the multitude of variables (over 138), and the dynamic and evolving nature of these markets, we adopt an exploratory approach for this objective. This approach aids in identifying and understanding the significant determinants in the price formation of Bitcoin and Altcoins, freeing us from the constraints of predefined hypotheses.

While a hypothesis-driven approach can provide specific directional expectations, our exploratory method ensures a comprehensive and unbiased discovery of key factors that influence the price formation of Bitcoin and Altcoins. It further allows us to examine how the impact of these determinants fluctuates across different market periods, mirroring the evolving nature of cryptocurrency markets.

For the second research question, our focus shifts towards understanding how these factors' influences change across varying market periods, further highlighting the dynamic nature of these markets. By taking this exploratory route, we aim to encompass the broad range of potential influencing factors without the confines of pre-established hypotheses, providing a more comprehensive understanding of these markets.

Objective 3: *To identify the factors responsible for price inconsistencies across different major cryptocurrency exchanges.*

- **Research Question 3.1:** What factors contribute to price inconsistencies across major cryptocurrency exchanges (Binance, Kraken, Coinbase)?
- **Research Question 3.2:** How does the influence of these factors vary across different cryptocurrency exchanges (Binance, Kraken, Coinbase)?

We maintain our exploratory approach for Objective 3 based on the rationale established for Objective 2. This exploratory approach is particularly beneficial when addressing the dynamic nature of these marketplaces and the multitude of possible influencing factors. In alignment with the exploratory nature of our research questions, this non-hypothesis-driven exploratory approach allow us to discover influential factors exhaustively and impartially without the constraints of a predetermined hypothesis.

Objective 4: *To compare the liquidity of cryptocurrency exchanges with that of different sizes of selected stocks using the Martin Liquidity Index (MLI).*

- **Research Question 4.1:** How does the liquidity of traditional stock exchanges (NYSE, NASDAQ, BSE Sensex, NIFTY) compare to that of major cryptocurrency exchanges (Binance, Kraken, Coinbase) as measured by the Martin Liquidity Index (MLI)?
    - **H1**: The liquidity of traditional stock exchanges (NYSE, NASDAQ, BSE Sensex, NIFTY) is significantly higher than that of major cryptocurrency exchanges (Binance, Kraken, Coinbase) as measured by the Martin Liquidity Index (MLI).

- **Research Question 4.2:** Which of the major cryptocurrency exchanges (Binance, Kraken, Coinbase) exhibits the highest liquidity as measured by the Martin Liquidity Index (MLI)?

This question is inherently comparative, aiming to rank the cryptocurrency exchanges based on their respective liquidity levels. As such, it does not seek to confirm or refute a predicted outcome, which is the typical role of a hypothesis. Instead, we directly examine and compare the MLI values of the three exchanges, providing an unbiased ranking of liquidity performance. As such, we proceed with our analysis without a prespecified hypothesis, thereby staying aligned with the exploratory and comparative nature of this research question.

Each of these hypotheses will be tested using relevant data and methodologies during the research. The Altcoins considered for the scope of this study are Ethereum (ETH), Ripple, Dash, and Litecoin.

1.7 Scope of the work

This research is focused on investigating the dynamics of pricing behavior and liquidity in the cryptocurrency market, concentrating primarily on Bitcoin and selected Altcoins (Ethereum, Ripple, Dash, and Litecoin). Our investigation is organized into four principal sections:

- **Cryptocurrency Price Trends and Market Evolution**: This section of the research investigates the trends in exchange prices for Bitcoin and Altcoins with an intention to predict future exchange prices of Bitcoin and the Altcoins in consideration. It considers the statistical characteristics of these price trends and evaluates their transformation across different market periods.
- **Factors Influencing the Price of Cryptocurrencies**: This segment identifies the crucial factors that impact the price formation of Bitcoin and Altcoins, and how these influences fluctuate across different market periods.
- **Price Inconsistencies in Cryptocurrency Pricing Across Different Exchanges**: This part of the research aims to determine the reasons for price inconsistencies across major cryptocurrency exchanges such as Binance, Kraken, and Coinbase. Here, we evaluate what factors influence the prices in different cryptocurrency exchanges.
- **Liquidity Analysis of Cryptocurrency and Traditional Stock Exchanges**: The last part of this work compares the liquidity of cryptocurrency exchanges with the liquidity of different sizes of selected stocks using the Martin Liquidity Index (MLI).

1.8 Thesis Structure and Outline

This thesis begins with an **introductory chapter** laying the research's foundation. It explains the background and rationale of the study, clearly stating the objectives that guide the investigation. This chapter also addresses the research questions and proposes hypotheses that provide a roadmap for the research. The section also defines the scope and outlines the limitations, setting the boundaries within which the research is conducted.

Chapter 2 presents a comprehensive literature review that broadens the research context. It opens with an overview of cryptocurrencies, thoroughly understanding their origins, workings, and significance. Following this, the focus shifts to the dynamics of cryptocurrency markets. The literature review then proceeds to investigate the role of cryptocurrency exchanges, their operation, and how they contribute to the larger cryptocurrency ecosystem. It concludes with an exploration of liquidity in cryptocurrency markets, a vital aspect in determining the efficiency and viability of these markets.

Chapter 3 elaborates on the research methodology employed in the study. It outlines the distinct research methodologies for each of the four objectives, addressing the rationale for selecting these methodologies. This chapter also elaborates on model specification, assumptions, and the procedure for analysis for each objective. Furthermore, the data collection process, including sources and collection methods for each objective, is meticulously explained. The chapter concludes with details of the data analysis and pre-processing techniques, evaluation metrics, and a discussion on the limitations and ethical considerations of the research.

Chapter 4 presents the results and discussion related to each research objective. It provides an overview of the results, interprets them, and compares these findings with baseline data and existing literature. It takes the reader through a comprehensive evaluation of the trends and forecasts of Bitcoin and Altcoin prices, price formation factors, factors for price inconsistencies across exchanges, and liquidity comparison.

It begins with studying trends and forecasts of Bitcoin and Altcoin prices, interpreting how these findings interweave with the broader fabric of cryptocurrency economics. This is followed by an investigation and discussion of the factors affecting the price dynamics of Bitcoin and Altcoins. The chapter then addresses price inconsistencies across various cryptocurrency exchanges, illuminating the unique attributes of platforms like Binance, Kraken, and Coinbase. Lastly, the chapter investigates liquidity across major cryptocurrency and traditional stock exchanges, juxtaposing findings to create a cohesive view of liquidity dynamics. Each segment of this chapter offers valuable insights, paving the way for a comprehensive understanding of the intricacies of cryptocurrency pricing and liquidity.

Chapter 5 concludes the research and contains a comprehensive summary of the findings from the research and discusses their implications. This chapter begins with a summary of findings for each objective, providing an overarching view of the research outcome. The research implications are then discussed in terms of advancing theoretical understanding and its utility for practitioners and policymakers. This chapter also acknowledges the current study's limitations and suggests areas for future research. The chapter concludes with a synthesis of the key findings from the study.

Chapter 2

Literature Review

Satoshi Nakamoto, a pseudonymous individual, or group, introduced cryptocurrencies by creating Bitcoin, the first peer-to-peer cryptocurrency (Nakamoto, 2008). Bitcoin was initially developed as a decentralized, peer-to-peer electronic medium of exchange for digital transactions using Blockchain as the underlying technology(Stancel, 2015). Blockchain is a distributed ledger system that ensures transaction transparency, security, and immutability of digital transactions over a computer network without a central authority. Bitcoin's introduction, use, and widespread adoption revolutionized the financial industry by providing an alternative medium of exchange without the need for intermediaries such as banks and by facilitating direct peer-to-peer transactions. Bitcoin was first widely used as a form of peer-to-peer trading due to its minimal transaction fees and fast confirmation time (Schueffel, 2021). However, as Bitcoin's popularity increased, it began to be used as a speculative asset. Several alternative cryptocurrencies, known as Altcoins, were created due to Bitcoin's rising prominence. These Altcoins offered distinctive characteristics and use cases, substantially expanding the cryptocurrency landscape (Ong et al., 2015). Bitcoin, which is designed to have a maximum limit of 21 million units that can ever exist, is immune to conventional inflation that can erode the value of traditional currencies. This inherent feature showcases the unique dual role of Bitcoin as both a medium of exchange and a store of value. Therefore, it presents an appealing option for investment, given its potential to increase in value over time (Sauer, 2016). (Salman & Razzaq, 2018) further highlighted the transformative impact of blockchain and Bitcoin on the financial sector, suggesting that the decentralized nature of blockchain could redefine traditional financial systems, reinforcing Bitcoin's significance as an investment mechanism in the evolving financial landscape.

2.1 Cryptocurrency Price Trends and Market Evolution

The volatility of Bitcoin's price has been a focal point for individual and institutional investors since its inception. Bitcoin and other cryptocurrencies offer high-risk, high-return opportunities to investors. However, the inherent volatility of their prices poses a substantial risk of capital loss for investors. According to the data accessed from www.coinmarketcap.com on July 15, 2023, Bitcoin has demonstrated significant price swings across daily, weekly, monthly, quarterly, and yearly scales. Shifts in cryptocurrency market behavior have been observed over time, often coinciding with major events or trends. These shifts can be precipitated by factors such as technological advancements, regulatory changes, market maturation, or significant market events. Identifying and understanding these shifts in price trends is crucial for investors, market participants, and researchers who seek to develop improved price forecasting models. Traditionally, Bitcoin and Ethereum, due to their largest market capitalization among all cryptocurrencies, have been closely associated with influencing the market sentiment of the entire crypto sector.

According to the previous literature, the price history of Bitcoin and other cryptocurrencies is often partitioned into distinct eras, each delineated by significant events and patterns. In 2014, Mount Gox, a major Japanese exchange, was hacked, resulting in thousands of investors losing their Bitcoin assets. This crucial event in the world of cryptocurrencies prompted the Japanese government to introduce new regulations for cryptocurrencies, a response that garnered international attention (Ishikawa, 2017). The Mount Gox incident not only marked a turning point in the perception of the security of digital assets but also highlighted the critical need for regulatory intervention in the nascent cryptocurrency industry. In the cryptocurrency literature, the Pre-Mount Gox crash period is characterized by a particular set of price swings, which evolved following the crash until the 2017 price boom. The period stretching from 2017 to 2021 represents another distinctive phase in cryptocurrency history, defined not only by its increased mainstream adoption but also by intensified regulatory scrutiny (Feinstein & Werbach, 2020). These developments significantly influenced price movements, adding a new layer of complexity to the cryptocurrency market dynamics. The current phase, post-2021, represents a novel set of trends and dynamics as the cryptocurrency market continues to mature, reflecting the ongoing evolution in this financial sphere. Figure 1 illustrates the market price trajectories of Bitcoin and Ethereum, the dominant cryptocurrencies that together represent around 70% of the overall market capitalization as of July, 2023.

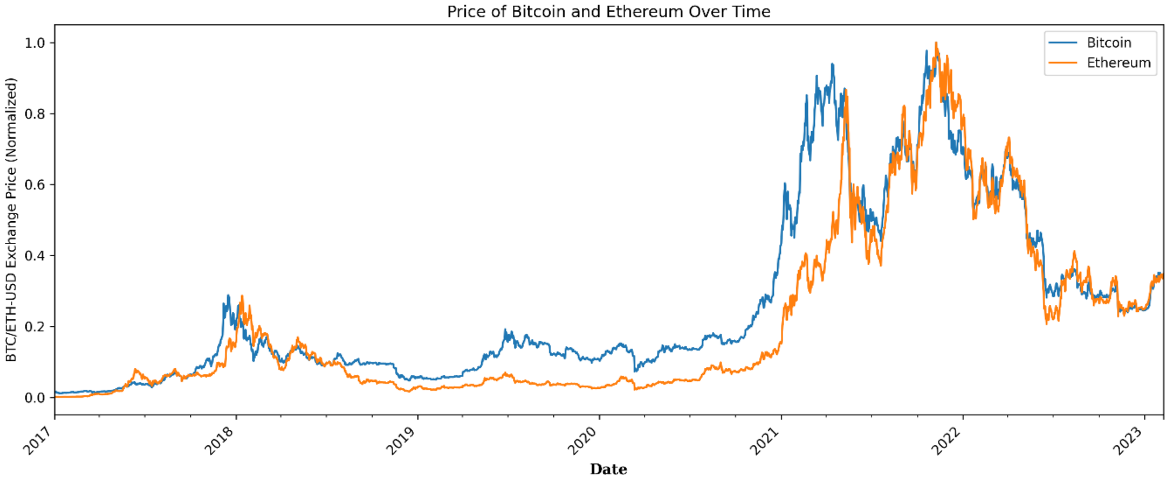


Figure 1. Historical Closing Prices of Bitcoin and Ethereum, the Two Leading Coins in the Cryptocurrency Space (Normalized for Comparison)

2.1.1 Pre-Mount Gox Crash Period history of price swings

In the period leading up to the infamous Mount Gox crash in 2014, the Bitcoin got excessive media attention due to its price fluctuations, the utility it offered and the other use cases associated to Blockchain technology. During this period, the price of Bitcoin significantly increased from a few cents to over $1,000. This increased investor interest and enthusiasm about the potential of the technology, attracting entrepreneurs. Early academic research identified that Bitcoin's average volatility outstripped traditional assets like gold and global currencies (Yermack, 2015) . However, this phase also exposed the market's vulnerability to manipulation, fraud, and security breaches, as evidenced by the Mount Gox incident. The crash catalyzed regulatory changes, with Japan introducing its first regulations addressing virtual currencies (Ishikawa, 2017). Furthermore, the period was characterized by the emergence of notable price bubbles, leading to substantial market volatility and eventual bursts, as exemplified by the collapse of the Mt. Gox exchange (Dwyer, 2015).

2.1.2 Price Trends Post Mount Gox Crash and Pre-2017

The market phase between the Mount Gox crash and 2017 was characterized by intensifying regulatory scrutiny, maturing market dynamics, and a slow recovery in prices (Ishikawa, 2017). Bitcoin, along with other cryptocurrencies, began to garner more mainstream attention, while efforts to bolster security and transparency in the market amplified. Despite persistent volatility, cryptocurrencies gained increasing recognition as a novel asset class during this period (Corbet et al., 2017). Nonetheless, price movements remained largely erratic, punctuated by occasional drastic price swings.

2.1.3 Price Trends between 2017 to 2021

From 2017 to 2021, the cryptocurrency market experienced an unprecedented level of volatility. In the final quarter of 2017, Bitcoin and other Altcoins driven by speculation, underwent an unprecedented bull run, peaking at nearly $20,000 in December before witnessing a sharp decline in the subsequent year. Studies from this period revealed strong bubble signals in Bitcoin starting from May-September 2017, reaching a critical point in mid-December 2017. In the same vein, Ethereum presented bubble signals in mid-June 2017 with a second, albeit weaker signal, around mid-January 2018, aligning with significant market crashes (Bianchetti et al., 2017). The period saw a surge in participation from institutional investors, enhanced regulatory intervention, and a diversifying array of cryptocurrency options. Despite ongoing price volatility, integrating Blockchain technology across various industries demonstrated the increasing mainstream acceptability of cryptocurrencies among retail and institutional investors (Liew & Hewlett, 2017). Individual investors have installed cryptocurrency wallets and initiated active trading to capitalize on the potential for high returns. In addition, institutional investors began allocating a portion of their portfolios to cryptocurrencies in recognition of the diversification potential (Nosratabadi et al., 2020). These new cryptocurrencies promised unique features such as enhanced privacy or faster transactions, however, most of the new Altcoins lacked practical business use cases or technological solutions, serving as speculative investments rather than innovative solutions (Kuo Chuen et al., 2017).

2.1.4 Post 2021 Period history of price swings

Beyond 2021, the cryptocurrency market continues to exhibit an evolving landscape, shaped by aspects such as technological innovations, regulatory changes, macroeconomic factors, and shifting investor sentiment. Although price volatility endures as a market constant, the rise in institutional investor involvement, continuous development of cryptocurrency infrastructure, and increasing use of cryptocurrencies within the wider financial ecosystem have collectively resulted in constantly changing market dynamics (Tripathi & Sharma, 2022). Bitcoin reached its all-time high price of $64,863 in November 2021 (Younis Masiha, 2022). However, Bitcoin's price declined to approximately $36,654 by January 2022, marking a nearly 43% decrease in less than three months. In contrast, Bitcoin's price increased from about $30,681 to $59,519 within a single day in May 2021, an extraordinary daily price increase of around 94%.

2.1.5 Past Works

The observed price fluctuations in Bitcoin prices suggest its dual nature. It presents an opportunity for considerable returns, reflecting the advantageous aspect of its decentralized and dynamic market driven by supply and demand (Goczek & Skliarov, 2019). At the same time, presents considerable risks to traders and investors and challenges researchers in formulating precise price prediction methodologies. Moreover, the dynamic nature of the cryptocurrency market, characterized by evolving market cycles, changes in the public interest, advancements in Blockchain technology, and varying degrees of speculation and adoption, further complicates price prediction. The determinants influencing pricing in one phase of the market's evolution may not remain relevant in the subsequent phase, leading to investor uncertainty (Tripathi & Sharma, 2022). With its dynamic and evolving nature, the cryptocurrency market presents opportunities and challenges for investors and businesses, necessitating careful analysis of the price trends for informed decision-making. For researchers, the challenge lies in accurately predicting these price fluctuations. According to (Lahmiri & Bekiros, 2019a), the unpredictability inherent in the cryptocurrency market complicates forecasting efforts and necessitates the development of sophisticated predictive models. Additionally, due to their decentralized nature, cryptocurrencies are less likely to correlate with traditional investment assets, making them a complex asset to integrate into existing investment portfolios for institutions(Baur et al., 2018).

A considerable body of prior literature exists on cryptocurrency price forecasting and modeling. These studies can be broadly classified into three main groups based on the type of price modeling methodology: studies using Econometric and Statistical models, Machine Learning models, and Deep Learning models.

The preliminary studies applied traditional econometric and statistical models, to forecast the prices of Bitcoin and other cryptocurrencies (Cermak, 2017; Cheung et al., 2015; Ji et al., 2018; Kristoufek, 2015; Munim et al., 2019; Troster et al., 2019). While these approaches were helpful to improve the price prediction performance in some cases, but these methods faced challenges during the data periods of high-boom-bust cycles due to inherent assumptions that were often inconsistent with the complexities and volatility observed in the cryptocurrency market (Gyamerah, 2020; Hu et al., 2020). They relied on the principle of Gaussianity, which suggests the residuals, or errors, should adhere to a normal distribution. In contrast, cryptocurrency price fluctuations frequently displayed fat-tailed and skewed distributions, indicating a greater likelihood of extreme price variations than predicted by a normal distribution (Eom et al., 2019). Furthermore, the standard assumption of homoscedasticity (Jarque & Bera, 1980), implying consistent variance across all levels of independent variables, did not capture the typical heteroscedastic behavior of cryptocurrencies (Szetela et al., 2016). These models also presumed stationarity, implying that the properties of a time series remain consistent over time, an assumption that was unable to account for the dynamic and evolving nature of cryptocurrency markets (Balcilar et al., 2017a). The presumption of no autocorrelation between error terms was another limitation, as significant autocorrelation is often observed in cryptocurrency markets (Tartakovsky, 2020).

The next category of studies apply Machine Learning (ML) models, particularly those leveraging supervised learning techniques. These models - encompassing methods such as Support Vector Regression (SVR), Regression Trees, Random Forests, gradient boosting methods like LightGBM, XGBoost, and ensemble models offered improved forecasting performance (Z. Chen et al., 2020; Jang & Lee, 2018a; Lahmiri & Bekiros, 2020). Such ML models demonstrate proficiency in decoding complex patterns when either integrated by traditional statistical or econometric methods with strong data pre-processing approaches (Gyamerah, 2020). Nevertheless, even with their strengths, they face unique challenges. Overfitting is a common issue with most machine learning models in which the model excels in the training data but performs poorly with unseen data. Furthermore, the performance of these supervised learning models tends to degrade with high-dimensional, multivariate input feature sets, indicating their susceptibility to the Bellman’s curse of dimensionality.

Finally, the third category comprises deep learning models. These models capitalize on the benefits of neural networks and are particularly adept at capturing complex patterns in large datasets. Their capacity to model non-linear relationships and learn hierarchical representations has shown promise in the chaotic and complex domain of cryptocurrencies.

The available literature on deep learning models for cryptocurrency price forecasting, returns forecasting, mid-price forecasting and other price trend modeling objective is substantial. While our research does not encompass all existing studies, we aim to present an overview of key research pieces, acknowledging their limitations and untouched areas. Our primary goal is to propose an advanced forecasting architecture that addresses the limitations of existing research and improves the effectiveness of deep learning in financial time series forecasting. Through this, we seek to bridge the gaps identified in the academic discourse and contribute to the evolution of the field using novel methodologies not extensively explored before. Table 1 presents a chronological review of important previous studies in Cryptocurrency price prediction. This helps to contextualize our research by situating it within the broader academic discourse, shedding light on the evolution of forecasting methodologies over time, and indicating the progression of thought and techniques in this field. We form the foundation for our proposed advanced forecasting architecture and develop our innovative methodology by identifying the limitations and gaps of methods and data used in these studies.

Table 1 A chronological review of past works on cryptocurrency price forecasting.

Study	**Data Frequency and Duration**	**Methods**	**Dependent Variable**	**Independent variables**	**Main Findings**	**Research Gap(s)**
(Ranjan et al., 2023)	Daily, 5-min intervals November 2, 2016, to June 11, 2018	Random Forest, XGBoost, Quadratic Discriminant Analysis, Support Vector Machine, Decision Tree, K-Nearest Neighbors	Bitcoin Close price	Market capitalization, Block size, Hash rate, the Bitcoin code, Mining difficulty, Time between blocks, Trades per minute, Number of transactions, Confirmed transactions per Day, Mempool transaction count, Estimated transaction value, Total transaction fees, Mempool size, Google trend search volume index, gold spot price	The paper concludes that machine learning algorithms can be used effectively for predicting Bitcoin prices, and that simple statistical models can be used for high-dimensional features with complex models for low-dimensional features, in order to avoid overfitting.	Features like media speculation and sentiment on social networks not considered
(Rathee et al., 2023)	Daily data from 2015 to 2021	MLP, GRU, CNN and LSTM, Ensemble-based neural networks with CNN-BiLSTM for long-term price prediction of cryptocurrencies	Close prices of Bitcoin, Ethereum, Dogecoin and Litecoin	Open, High, Low, Close, RSI	CNN-BiLSTM ensemble model is the best performing model	1. Time period (Date, month, year) in the dataset not described correctly. Only year has been described. No Train and test periods 2. Hyperparameter optimization method not specified. 3. Only one interval used to generalize the findings of the results.
(Yang et al., 2023)	Experiment 1 and 2: 11 Sept. 2022 to 25 Sept. 2022. Experiment 3: 11 Sept, 2022 to 20 Sept, 2022 Experiment 4: 13 April to 27 April, 2022 Daily data	Fractional Grey Model (FGM), Grey Model (GM), Particle Swarm Optimization	Bitcoin, Ethereum, and Litecoin closing price	Closing Price	The FGM (1,1) outperformed the GM (1,1)	1) Limited Time Scope: Only 15-day daily time period. 2) No External Factors: Only closing prices considered; no external factors used. 3) Stable Market Assumption: Assumes stable market conditions, ignoring inherent

						volatility of cryptocurrencies.
(Zhu et al., 2023)	January 4th 2023 to April 15th 2023	Support Vector Regression (SVR), Least Squares Support Vector Regression (LSSVR), and Twin Support Vector Regression (TWSVR) for prediction, Whale Optimization Algorithm (WOA) and Particle Swarm Optimization (PSO) for optimization of mixed models.	Bitcoin Closing price	27 variables as independent variables, divided into three secondary indices: Bitcoin-supply factors, Bitcoin-demand factors, and other cryptocurrency's prices.	The XGBoost-WOA-TWSVR combined model showcased the highest prediction effectiveness with an EVS score of 0.9547. Of the three prediction models, LSSVR was the quickest, followed by TWSVR and SVR. The WOA-SVR model also attained a substantial EVS value of up to 0.9491.	Benchmarking limited to the ARIMA model, potentially providing an incomplete comparison. No juxtaposition with past state-of-the-art methods, limiting evaluation context. Model efficiency assessment based solely on CPU time, possibly overlooking other efficiency aspects. The chosen models (SVR, LSSVR, TWSVR) demonstrated difficulty in handling data with large fluctuations.
(G. Kim et al., 2022)	March 27, 2018 to November 16, 2021; Daily data	Data normalization and segmentation using PELT analysis, self-attention-based LSTM with Change point detection	Bitcoin, Ethereum, Litecoin and Ripple Closing price	254 variables categorized into multiple groups namely: price, adoption, market, and valuation	SAM-LSTM with Change point Detection is the best performing model with MAE, RMSE, MSE, and MAPE values of 0.3462, 0.5035, 0.2536, and 1.3251, respectively	The study uses only Blockchain (on-chain) data; omits other potential influential factors. Results validated via univariate BTC price prediction, may not apply to other cryptocurrencies.
(Maciel et al., 2022a)	January 1, 2018 to February 28, 2022 Daily data divided in 3 separated intervals	Level-set methods (LSM) m ARIMA, Multi-Layer Perceptron	Closing prices of: Cardano, Binance Coin, Bitcoin, Ethereum, Chainlink, Litecoin, Tron, Stellar, Monero, and Ripple	Lagged values of the series for both Autoregressive Integrated Moving Average and Multilayer Perceptron models.	Level Set Method applied for crypto price prediction, outperforming Random Walk and Autoregressive Integrated Moving Average models. Level Set Method found in 83% of superior model sets, Autoregressive Integrated Moving Average in 32%. Test statistics for Level Set Method found to be higher than Random Walk model.	1. No comparison to strong algorithms, only linear/nonlinear methods used. 2. Market sentiment impacts not assessed. 3. Macroeconomic influences on prices overlooked. 4. Impact of other digital assets unaccounted for.
(Luo et al., 2022)	November 24th, 2017 to November 27th, 2020	Empirical Mode Decomposition, Variational Mode Decomposition, ARIMA, GARCH, SVR, Backpropagation neural network RNN, KNN, Gradient Boosting Decision Tree,	Bitcoin Closing price	Trading information, Hyper frequency price, Supply and demand factors, Energy consumption of Bitcoin mining	Study introduced a multiscale analysis and deep learning model for accurate Bitcoin price prediction. Bitcoin price series components (high, medium, low-frequency) identified using EMD and VMD.	1. Study overlooked the influence of key factors like trading volume and macroeconomic indicators on cryptocurrency prices. 2. Multiple input

		Random Forest, XGBoost, LSTM, and Extreme Learning Machine.			LSTM and ELM employed for high and low-frequency data predictions. LSTM outperformed ARIMA and SVR benchmark models in prediction accuracy.	features like trading information, hyper frequency price, supply-demand factors, and Bitcoin mining energy consumption were not integrated into multiscale analysis and deep learning.
(Tanwar et al., 2021a)	Litecoin: August 24,2016 to May 26,2021 Zcash: October 29, 2016 to May 26, 2021	LSTM, Grated Recurrent Unit(GRU), Hybrid model	Litecoin and Zcash Closing price	1D,3D,7D,30 Days Closing price for Litecoin and Zcash	1. The study compares a hybrid Deep Learning model to LSTM and GRU for cryptocurrency forecasting. 2. The hybrid model consistently shows lower MSE losses across all window sizes than the LSTM and GRU models.	1. No consideration given to other cryptocurrencies and external factors affecting pricing. 2. Only Mean Squared Error (MSE) is used as the evaluation metric which is an insufficient assessment of model's performance.
(Miller & Kim, 2021)	Start dates: Bitcoin, Ethereum, Ripple, Tether, Dogecoin from January 1, 2017; Binance Coin from August 15, 2017; Cardano from October 10, 2017; USD Coin from October 11, 2018; Uniswap from September 18, 2020; and FLOW from January 29, 2021; End date for all cryptocurrencies: July 14, 2021.	Study used statistical and machine learning methods, including RNNs, DNNs, LSTMs, Holt's exponential smoothing, autoregressive integrated moving average, and ForecastX to predict log returns.	price log returns of Bitcoin, Ethereum, Tether, Ripple, Binance Coin, Cardano, FLOW, USD Coin, Dogecoin, Uniswap	log returns of the top 10 cryptocurrencies	Multivariate LSTM outperformed univariate time-series methods for predicting top 10 cryptocurrency prices	1. Assumption of linearity: The formulation assumes that the price log return of a given cryptocurrency is a linear combination of the price log returns of all cryptocurrencies at the previous time step. 2. No comparison with past State-of-the-Art.
(W. Chen et al., 2021)	Period 1: August 1, 2011 to December 31, 2013 Period 2: August 1, 2013 to December 31, 2014 Period 3: July 1, 2014 to December 31, 2017 Period 4: July 1, 2015 to July 31, 2018	Random Forest and ANN used for feature selection. ANFIS and GA employed for dimensionality reduction and parameter tuning. Back-propagation and least squares for refining membership function parameters. Gaussian function as the fuzzy membership function. GA matches trials with parameter values.	Bitcoin exchange rate	Nine Blockchain information variables; Public Attention: Tweets, Google Trends; Economic Factors:8 macroeconomic variables; five Global Currency to US dollar Ratios	The study predicted Bitcoin exchange rate using economic and technology determinants. ANN and RF were used for feature selection. Predictors were integrated into LSTM for predictions. Performance of LSTM was better than ANFIS, SVR, ARIMA, and previous rate-based LSTM	The study was exhaustive but data period is outdated and may not be a representative of more recent trends or changes in Bitcoin exchange rate determinants.

(Shahbazi & Byun, 2021)	Litecoin: Data from 2016 to 2020 with 1276 data points Monero: Data from 2015 to 2020 with 1276 data points	The study used Reinforcement Learning (RL) for Litecoin and Monero price prediction.	Next day, 3 day, 7 day and 30 day exchange price for Litecoin and Monero	Blockchain information, OHLC prices, Volume	The proposed Reinforcement Learning (RL) approach integrated with blockchain for Litecoin and Monero price prediction achieved superior performance compared to other state-of-the-art methods.	The study only focuses on Litecoin and Monero. Model generalizability to major cryptocurrencies such as Bitcoin and Ethereum is not tested.
(Gao & Su, 2020)	June 2, 2016 to December 30, 2018 Daily data	BP Neural Network, Least Square-SVM, Particle Swarm Optimization	rate of return of Bitcoin and Ethereum	Bitcoin and Ethereum Close price time series	PSO-LS-SVR shows best result fit in Bitcoin and Ethereum prediction compared to BP Neural Network and SVR algorithm. SVR algorithm exhibits better learning, model fitting, and prediction than BP Neural Network.	No external input factors used. Study uses univariate data.
(Mallqui & Fernandes, 2019)	2 intervals Interval 1: August 19th, 2013 to July 19th, 2016 Interval 2: April 1st, 2013 to April 1st, 2017	SVM algorithm, ANN, RNN, Machine Learning Ensemble algorithms (k-Means, RNN, Decision Tree)	maximum, minimum, and closing Bitcoin exchange rates of Bitcoin	variables grouped into: 'Day D', 'Day (D - i)', '30-day WMA', and various market indicators, including trade volumes, transaction metrics, hash rate average, and closing prices of key indexes.	Relief technique stood out for attribute selection; SVM algorithm excelled in all price predictions (max, min, closing) for both intervals.	Accuracy of maximum, minimum, and closing price predictions for Bitcoin is limited, given the high MAE and MAPE values obtained.

Additionally, the performance of alternative cryptocurrencies, or 'Altcoins', can also exert an influence on the overall market, indicative of shifting risk tolerances and investor sentiment (Hayes, 2017). As per the data retrieved from www.coinmarketcap.com on July 15, 2023, more than 10,000 altcoins are listed, with Ethereum, Ripple, Litecoin, Dash, and Cardano standing out due to their popularity and widespread usage. These Altcoins have gained significant attention and investment due to their unique features and potential for growth. Altcoins provide various innovative solutions within the cryptocurrency ecosystem, addressing unique use cases. For instance, Ethereum, launched in 2015, is a Turing-complete programming platform where each node activates and executes blockchain contracts utilizing Ether, its indigenous cryptocurrency. Ethereum's market dominance is largely attributable to its smart contract functionality, a feature that has attracted the attention of developers and institutional entities and contributed to its market dominance (Mohanty, 2018). Similarly, Dash coin provides an interesting use case in which its unique distribution system incentivizes miners and controller nodes, thereby contributing to the security and functionality of the network. Furthermore, a fraction of rewards is set aside for the treasury, supporting continual development, community projects, and adoption-promoting marketing strategies (Biryukov & Tikhomirov, 2019). Ripple's XRP, which operates on a "global consensus ledger" rather than a typical blockchain, serves the specific needs of the banking industry (Akcora et al., 2022). Litecoin, by employing the Scrypt mining algorithm instead of Bitcoin's SHA-256, further exemplifies the variety and specialization within the digital asset landscape.

This diversification of digital assets, facilitated by their unique use case solutions, caters to different sectors, and attracts a wider range of retail and institutional investors, contributing to the growth of the cryptocurrency market. Nonetheless, this diversity in the landscape of digital assets increases volatility and speculation risks. Numerous newly issued coins are primarily motivated by speculative motives rather than fundamental utility or technological innovation, which can cause market instability necessitating careful price trend analysis for investing (A. D. Lee et al., 2020). Another challenge to the accurate price forecasting of cryptocurrencies is volatility fluctuation. Cryptocurrencies can display diverse volatility profiles based on factors such as market capitalization, liquidity, their inherent technological attributes, and the evolving nature of the market. Some cryptocurrencies might exhibit persistent volatility, while others may experience episodic spikes (Balcilar et al., 2017a; Liang et al., 2020; Woebbeking, 2021). These diverse volatility patterns call for advanced models capable of capturing such heterogeneity in volatility structures. This dynamism is not only tied to the market's rapid growth and evolution but also to the continuous influx of new participants, technological advancements, and regulatory changes. Consequently, the statistical attributes associated with cryptocurrencies, such as their mean, variance, and higher-order moments, may display substantial temporal variations (Eom et al., 2019; Takaishi, 2018). This phenomenon mandates consistent reassessment of the models used for price prediction and necessitates model flexibility to capture the shifting dynamics accurately.

Given the observed complexities in cryptocurrency price series, such as non-linearity, multi-fractality, leptokurtosis, and skewness, there is a necessity for non-linear analysis-based methods for accurate price prediction of Bitcoins and other cryptocurrencies (Bouoiyour & Selmi, 2017; Lahmiri & Bekiros, 2019a; Lamothe-Fernández et al., 2020) . Deep Learning (DL) models have emerged as promising tools in this context. Models such as ANN (Soni, 2012), LSTM (Luo et al., 2022), BiLSTM (Massaoudi et al., 2020), GRNN (Abhyankar & Singla, 2022), and CNN (Xin & Peng, 2020) among many others have been successfully applied in past studies to improve the forecasting performance in non-linear datasets. By leveraging multiple layers of artificial neurons and non-linear activation functions, DL models can capture complex, non-linear relationships without relying on the restrictive assumptions of traditional econometric models (Lamothe-Fernández et al., 2020). However, one important issue with existing studies that use deep learning is the interpretability of the DL model. This lack of explainability can hinder the understanding of crucial aspects driving price changes and can also reduce trust in the model's predictions, particularly in a domain as sensitive as financial forecasting. The interpretability of the Deep Learning models Cryptocurrency price forecasting can be provided through widely accepted techniques like Shapley analysis (Ning et al., 2022). This research gap presents an opportunity to investigate the application of explainable AI techniques in cryptocurrency price forecasting. By integrating interpretability methods into DL models, we can gain insights into the factors driving cryptocurrency prices and enhance our understanding of this volatile market. Furthermore, the integration of explainable AI in these DL-based models can clarify the decision-making processes for enhancing trust and understanding of results.

Prior studies have applied a variety of fundamental and technical indicators as input predictor variables in cryptocurrency price forecasting. Broadly, these variables can be segmented into two categories: fundamental and technical indicators. Fundamental Indicators consist of intrinsic factors that can impact the value of an asset. Notable among these factors are blockchain metrics like Hash Rate, Difficulty, Active Addresses, Mining Costs, and Whale Activity. Many studies have also used several other predictor variables such as Google Search Trends Data, Global currencies to USD ratio, Oil price, gold price, unemployment data, and sentiments from social media platforms. These indicators aim to measure the cryptocurrency's inherent value and understand its functional dynamics within a broader economic and regulatory landscape. Technical Indicators, on the other hand, are derived from statistical and mathematical assessments of past market data, predominantly price and volume. Examples include the Relative Strength Index (RSI), Moving Averages (e.g., 50-day and 200-day), Bollinger Bands, and MACD (Moving Average Convergence Divergence). These indicators primarily serve to identify market patterns and potential price trajectories based on historical data.

Furthermore, it has been observed that effective data pre-processing techniques, such as removing inherent signal noise (Xin & Peng, 2020), eliminating outliers (C. F. Lee & Lee, 2015), and employing smoothing techniques such as Kalman filters, Savitzky Golay filters and exponential smoothing (Massaoudi et al., 2020; Tang et al., 2021), can considerably improve the predictive accuracy of these models by reducing the noise complexity of data (Massaoudi et al., 2020; Seo et al., 2018). The application of feature selection techniques is also beneficial, emphasizing the need to concentrate on the most pertinent indicators (Kursa & Rudnicki, 2010; Radojičić et al., 2021; Sugunnasil & Somhom, 2010). However, a significant research gap emerges upon closer examination of the literature, pointing to the need for further exploration. While many studies have utilized either fundamental or technical indicators, few have integrated both sets of metrics into a comprehensive input dataset and subsequently developed a robust price forecasting model. Additionally, factors such as GitHub statistics, which provide insights into the developmental progress of a project, remain largely uncharted in past studies. A comprehensive study leveraging an exhaustive dataset combining both fundamental and technical analysis is markedly absent. Such a study would ideally employ robust data preprocessing methodologies using advanced signal processing techniques tailored to the non-linear nature of cryptocurrency data, and subsequently deploy deep learning techniques for forecasting across multiple intervals. Existing studies have yet to comprehensively explore a financial forecasting architecture that consolidates these advanced methods.

We summarize the key research gaps after carefully reviewing existing forecasting methodologies and identifying their limitations. These gaps reveal the shortcomings in current approaches and establish the areas where our thesis contributes significantly. We aim to address specific challenges in cryptocurrency price forecasting, paving the way for advancements in this critical study area.

2.1.5.1 Summary of Research Gaps for Objective 1

i. **Existing Limitations in Forecasting Methodologies**: Previous literature has predominantly utilized traditional econometric and statistical models, which often struggle during the high volatility and cyclicality endemic to cryptocurrency markets due to their reliance on linear assumptions. These models fail to adequately characterize cryptocurrency price trends due to their volatile, nonlinear, multifractal, multi-scaled, and dynamic nature. While machine learning and deep learning models offer some improvements in predictive capability, they frequently lack interpretability, an essential aspect of reliable forecasting in financial markets.

ii. **Integration of Fundamental and Technical Indicators**: Previous studies have typically focused on either fundamental or technical indicators but rarely integrated both within a comprehensive forecasting model. Additionally, the impact of variables such as media speculation, social network sentiment, and other macroeconomic factors has been underexplained in the literature. An approach combining fundamental and technical indicators along with broader economic and social factors may yield improved cryptocurrency price forecasting performance.

iii. **Model Generalizability and Data Scope**: Previous research has often been restricted in data scope, frequently constrained to limited periods or individual cryptocurrencies. The generalizability of models across different cryptocurrencies and shifting market conditions represents a substantial gap in the literature. Most existing studies fail to evaluate model performance across multiple cryptocurrencies using extensive historical data spanning various market regimes. Testing model robustness and external validity across diverse data is a critical area for improvement.

iv. **Advanced Preprocessing Techniques**: Previous studies have often overlooked critical data preprocessing techniques, which are vital for enhancing model predictive accuracy. Specialized signal processing methods tailored to handle the nonlinear dynamics of cryptocurrency data represent an unmet need. The application of noise filtering, feature extraction, and data transformation techniques could better prepare cryptocurrency price data for modeling and improve forecasting performance. Further research into preprocessing techniques capable of capturing the distinct characteristics of cryptocurrencies is necessary.

v. **Explainability in Deep Learning Models**: The existing literature shows a scarcity of studies incorporating explainable artificial intelligence techniques, such as SHAP analysis, into deep learning models for cryptocurrency price forecasting. This gap impedes understanding of the primary factors driving price dynamics. Incorporating methods to enhance model interpretability alongside predictive accuracy could bolster credibility and reveal novel insights into cryptocurrency behavior. Additional research is warranted to investigate explainable deep learning in the context of cryptocurrency forecasting.

Based on these identified gaps, Objective 1 of this thesis will focus on developing a robust and interpretable forecasting model that integrates both fundamental and technical indicators, employs advanced data preprocessing techniques, and ensures model explainability. This approach aims to overcome the limitations of current forecasting methodologies and provide a more accurate and reliable tool for understanding cryptocurrency price dynamics.

2.2 Price Formation Factors of Cryptocurrencies

Several factors contribute to cryptocurrency price formation. These include internal elements related to the Blockchain technology underlying the currency, such as its algorithmic structure, security measures, and governance protocols, as well as the extent of its adoption by users and vendors (Jang & Lee, 2018a). Furthermore, external factors such as regulatory shifts, fluctuations in market sentiment, technological progress, macroeconomic trends, and broader market sentiment significantly contribute to this volatility (Ciaian et al., 2016). (Tanwar et al., 2021a) argued that many technical, sentimental, and legal factors can affect the price of cryptocurrencies. Moreover, market speculation from diverse participants or even remarks from a prominent figure like Elon Musk, can substantially influence the cryptocurrency market (Bartoletti et al., 2021).

(Ciaian et al., 2016) found that market forces and Bitcoin attractiveness for investors and users have a significant impact on Bitcoin price but with variation over time. (Bouoiyour & Selmi, 2017) found that the use of Bitcoin in trade and the uncertainty surrounding China's deepening slowdown, Brexit and India's demonetization were found to be the most potential contributors of Bitcoin price when the market is improving. Lastly, the velocity of bitcoins in circulation, the gold price, the Venezuelan currency demonetization, and the hash rate were found to be the fundamental factors that influenced the Bitcoin price when the market is heading into decline. (Ciaian et al., 2018) investigated the price relations between Bitcoin and Altcoins and found that Bitcoin influences all Altcoins in the short term. However, neither similarity nor dissimilarity with Bitcoin could explain the long-term price relationships of Altcoins. The study found that macroeconomic indicators have a substantial impact on the prices of virtual currencies both in the short and long term. CNY/USD exchange rate was more impactful than USD/EUR in the long term, while the inverse was true in the short term. Shocks to gold and oil prices influenced most altcoins, but this was not statistically significant in the long term. Furthermore, Bitcoin prices significantly positively impacted Altcoin prices in the short term. In contrast, the currency supply had little impact on the prices of most virtual currencies.

(Carbó & Gorjon, 2022b) analyzed Bitcoin's price across three distinct time periods. Period 1: April 10, 2011 to August 30, 2013; Period 2: August 30, 2013 to December 19, 2017; and Period 3: December 19, 2017 to June 16, 2021. Each time period was characterized by a specific set of potentially explanatory factors, which were further classified into three distinctive groups namely technological variables, economic variables, and public attention variables. Initially, technological factors dominated price impact, contributing over 60%, but they decreased to 21% by the final period. Conversely, sentiment-related variables gained relevance, growing from 18% to over 40% in 2021. The authors used LSTM model for prediction and SHAP for results interpretation, they found Bitcoin price influencers evolved over time. In the first period, the most impactful features on Bitcoin's price were the SP500, the Nasdaq, the British Pound, the US Dollar, the Chinese Yuan, the Dax, the Nikkei, the Gold Price, the Oil Price, the VIX, the BTC Dominance, the number of Google searches, the number of Twitter followers, the number of Reddit subscribers, and the number of GitHub commits. By summing the Shapley values of each variable, the overall influence of the variable was determined. The analysis demonstrated that the British Pound exchange rate, the Nasdaq, and the Yuan were the most significant, each accounting for about 5% over the mean prediction. (Balcilar et al., 2017b) found that the relationship between Bitcoin returns and trading volume is not linear. In addition, this study revealed evidence of multiple shifts in the behavior of returns and their relationship to trading volume, indicating a dynamic and complex relationship.

(Hakim das Neves, 2020) examined the factors influencing Bitcoin’s price. These factors included macroeconomic and financial variables, attractiveness factors, and supply-demand dynamics. The authors found that while macroeconomic and financial elements have a short-term impact, they do not significantly influence Bitcoin's long-term pricing. Moreover, the price of gold was unrelated to Bitcoin's price, but the growing interest in Bitcoin significantly influenced its value. Furthermore, the research emphasized Bitcoin's price sensitivity to sudden global news events and crises, indicating the volatile nature of cryptocurrency prices. In contrast, (Gozbasi et al., 2021) found that gold prices do not significantly impact Bitcoin prices in the short or long term. Moreover, the study highlighted that the factors such as crude oil prices, the S&P 500 stock index, and volatility and financial stress indices consistently negatively impact Bitcoin prices, irrespective of the time frame. The research also unearthed a unidirectional causality from the volatility index to Bitcoin price and vice versa with the S&P 500 index, adding a renewed perspective to understanding price formation in cryptocurrencies.

According to (Meynkhard, 2020), the correlation between Bitcoin and Altcoin prices is flawless. This finding suggests that Bitcoin serves as a market driver in the cryptocurrency sector, as its price fluctuations impact the price fluctuations of specific Altcoins. Consequently, any factors affecting Bitcoin's price may indirectly impact Altcoin prices. This interdependence can be useful for predicting the price movements of associated Altcoins when Bitcoin's price fluctuates significantly.

(Tsang & Yang, 2020) demonstrated that the Bitcoin price is not simply determined by external influences, but also by internal factors such as price growth and transaction fee fluctuations. The fluctuations in Bitcoin transaction fees and price growth can create variances in Bitcoin prices across different exchanges, much like changes in Bitcoin's price could influence Altcoin prices. This indicates the presence of an interdependent relationship within Bitcoin's own market, where shifts in key internal factors can induce notable price dispersion. The results showed that during periods of high volatility, transaction fees and price growth were found to have a much larger influence on price dispersion, accounting for over 60% of the forecast error variance, as compared to around 20% during more stable periods. This finding implies that the effects of internal factors on Bitcoin prices are not static and can be magnified under specific market conditions. Therefore, one needs to consider these contextual variables when exploring Bitcoin price formation.

(H. Sabah, 2023) analyzed key drivers in Bitcoin price formation. The study found that factors like exchange rates, gold prices, stock market indices, and oil prices significantly affect Bitcoin prices. The influence of traditional stock market indices implies a strong correlation between traditional financial markets and Bitcoin. Moreover, the research highlighted the unique role of oil prices in Bitcoin's price formation, suggesting a potential risk-hedging mechanism for investors. Furthermore, the impact of gold prices and exchange rates underscores Bitcoin's perceived function as a 'digital gold' and its sensitivity to global economic conditions. This work offers crucial insights into Bitcoin's behavior within a broader financial ecosystem, aiding market participants to make informed decisions.

(Gaies et al., 2023) found that investor fear sentiment significantly influenced Bitcoin prices during the COVID-19 pandemic. Using a bootstrap rolling window Granger causality test, they discovered both positive and negative associations between fear sentiment and Bitcoin prices. The study showed that the interactions varied over time and were particularly prominent during periods of high volatility. These findings highlight the influence of investor sentiment on Bitcoin's price dynamics. (Kukacka & Kristoufek, 2023) demonstrated that a combination of fundamental elements, such as on-chain activity, macroeconomic factors, such as stock market price dynamics and uncertainty, and speculative attention-based factors significantly influences the formulation of Bitcoin's price. However, technical factors related to Bitcoin's supply, such as inflation and hash rate, did not significantly affect its price. The research also highlighted the importance of Bitcoin's production cost in explaining its price, thereby contesting the conventional economic view that Bitcoin lacks intrinsic value.

(Benlagha & Hemrit, 2023) found that used the Nonlinear Autoregressive Distributed Lag (NARDL) model to examine the relationship between Bitcoin's exchange price and various macro-financial variables. They observed significant short-term and long-term effects of inflation on Bitcoin's price, suggesting its potential as an inflation hedge. Moreover, their findings revealed a positive long-term correlation between market fear, measured by the Volatility Index (VIX), and Bitcoin's price, implying a possible refuge for investors during periods of increased market anxiety. Interestingly, they also discovered a consistently negative relationship between the Dow Jones Industrial Average (DJIA) and Bitcoin's price. Furthermore, they applied the quantity theory of money to explain how the size and velocity of the Bitcoin economy impact its price. Internet-based public interest, as tracked by Google Trends, was also found to be a significant determinant of Bitcoin's price. (Huang et al., 2019) conducted a study to understand the influence of technical indicators on the daily price movements of Bitcoin. The study used 124 technical indicators and a decision tree model to predict specific ranges of Bitcoin's daily price returns. The model outperformed the traditional buy-and-hold strategies. However, the study was limited to specific Bitcoin return ranges, suggesting further exploration of technical indicators' applicability in wider cryptocurrency price predictions and interactions with factors like market sentiment and regulatory changes.

(Ortu et al., 2022) explored the role of technical, trading, and social media indicators in predicting cryptocurrency prices, specifically Bitcoin and Ethereum. The dependent variable in the study was a binary representation of downward price movements. The study used a deep learning algorithm and found that augmenting traditional technical variables with social and trading indicators leads to improved classifications of price changes. The social media indicators were significant in price prediction, suggesting stakeholders should closely monitor crypto-related activities on social platforms to make informed decisions. However, the study was limited in its focus on only two cryptocurrencies and did not consider the impact of external factors, necessitating further research for comprehensive insights.

The above literature on cryptocurrency price formation highlights various influential factors, ranging from internal blockchain aspects to external market forces and sentiments. These encompass macroeconomic variables, attractiveness factors, market sentiments, and the varying technical, sentimental, and legal impact. While the studies offer valuable insights, they often employ machine learning or deep learning models for analysis, presenting an opacity challenge as these models operate as black boxes with limited explainability. Our work addresses this gap by applying Shapley analysis, offering a transparent and explainable model that elucidates the contribution of each feature to the prediction. A noticeable gap in the literature is the separate exploration of technical and fundamental indicators without considering their collective impact. Despite the individual significance of these indicators, an in-depth investigation of the combined effect on price formation has not yet been conducted in the existing literature. Additionally, the dependence of cryptocurrencies, especially Altcoins, on Bitcoin underscores the need for a holistic understanding of how factors affecting Bitcoin's price could indirectly impact Altcoin prices. Furthermore, the role of factors such as market sentiment, financial stress indices, gold prices, stock market indices, oil prices, and exchange rates on Bitcoin's pricing merits further exploration. Therefore, the contribution of our work lies in a comprehensive analysis of technical and fundamental indicators together and an emphasis on model explainability. This adds another dimension to understanding cryptocurrency price formation and assists market participants in making informed decisions.

2.2.1 Summary of Research Gaps for Objective 2

i. **Comprehensive Evaluation of Technical and Fundamental Variables**: The existing literature highlights various influential factors in cryptocurrency price formation, including internal blockchain aspects, external market forces, and sentiments. However, there is a paucity of research comprehensively evaluating the collective set of technical and fundamental indicators and their aggregated impact on price dynamics. An extensive investigation examining the effects of incorporating multiple classes of factors on predictive accuracy could provide vital insights into the underlying drivers of cryptocurrency fluctuations.

ii. **Need for Examining Interdependencies and Spillovers Between Cryptocurrencies**: Previous research indicates interdependence between cryptocurrencies, especially between Bitcoin and altcoins. However, further investigation is required to understand the subtleties of these relationships and implications on price dynamics. A more comprehensive inquiry into the interactions between major and minor cryptocurrencies is warranted to gain a deeper understanding of the underlying drivers of pricing behavior.

iii. **Role of Market Sentiment and Financial Stress on Cryptocurrency Pricing**: Prior studies have identified market sentiment and financial stress indices as important drivers of cryptocurrency price formation. However, their precise modes of impact and degree of influence warrant deeper investigation to further comprehension of their contributions to cryptocurrency pricing. Additional investigation of how factors such as on-chain activity, macroeconomic uncertainty, stock market fluctuations, and speculative attention affect cryptocurrencies warrants further research.

iv. **Impact of External Variables**: Prior research indicates that external factors, such as gold prices, stock indices, oil prices, and exchange rates, have noticeable impacts on cryptocurrency prices, especially Bitcoin. However, a thorough investigation into the precise nature and magnitude of their effects is still needed for a more complete understanding. Additional scrutiny on how major economic variables transmit external shocks to cryptocurrencies would support better pricing models.

v. **Enhancing Model Interpretability**: Many studies in the literature employ machine learning or deep learning models, which often lack model interpretability. Addressing this gap by utilizing methods such as Shapley analysis can offer transparent insights into the contribution of each feature to price prediction.

2.3 Factors Responsible for Price Inconsistencies across Cryptocurrency Exchanges

In economic theory, the Law of One Price (LOP) asserts that identical goods in an efficient market should have the same price across various locations, considering transportation costs and taxes (Miljkovic, 1999). It is an assumption that if there is a price difference, market participants will capitalize on this disparity until the price equalizes, a process known as arbitrage (Jouini, 2003). Cryptocurrency markets, operating uninterruptedly around the clock and throughout the week, leverage cutting-edge technology-based platforms and offer global accessibility, thus standing at the forefront of modern financial systems. (Pieters & Vivanco, 2017) demonstrated that the cryptocurrency market exhibits significant inconsistencies in its pricing across various cryptocurrency exchanges. According to (Zhang et al., 2018) cryptocurrency market is inefficient, which can result in significant price disparities between cryptocurrency exchanges. These disparities provide arbitrage opportunities for speculators to generate profits.

Past works show that price discrepancies in major cryptocurrency exchanges arise from several factors. These factors include the absence of a centralized regulatory authority, differing liquidity levels across exchanges, diverse market characteristics, and distinct regulatory practices across platforms (Aloosh & Ouzan, 2020; Cohen, 2021; Pieters & Vivanco, 2017). This complex landscape is further complicated as each exchange operates in a unique environment, hosting a distinct user base, security incidents and hacking of Blockchain ledger (Bartoletti et al., 2021) , fee structure, and potential market manipulation (Akba et al., 2021), among other factors. Therefore, the elements affecting price inconsistency might differ between exchanges, suggesting the utility of unique models for each platform to identify the specific factors contributing to price inconsistency effectively. (Pagnottoni & Dimpfl, 2019) analyzed the price discovery process for Bitcoin across different cryptocurrency trading platforms, utilizing a fundamental microstructure model. The findings suggested that Chinese cryptocurrency platforms, OKCoin and BTC-China, were primary contributors to Bitcoin's price discovery during the sample period. The inconsistencies in Bitcoin prices across exchanges were attributed to factors such as price fluctuations, variations in trading volumes, market hype, government regulations, the impact of exchange rates, and the dynamic nature of information sharing. However, the study suffered from limitations, including a narrow sample period, and a lack of consideration for exchange rate effects on Bitcoin prices. Despite these, the study's findings offer valuable historical insight, even though its immediate relevance may be reduced due to the Bitcoin ban in China post-2018 and other market developments since then.

(Makarov & Schoar, 2020) examined the arbitrage and price variations in Bitcoin markets across various global cryptocurrency exchanges. Their findings revealed that the price deviations were much larger across regions (such as the US, Japan, Korea, and Europe) than within the same country, indicating a segmented market. The underlying factors driving these price variations were differences in liquidity, exchange risk, and governance risk across regions. However, limitations such as reliance on linear models (factor analysis and canonical correlation), an insufficient exploration of arbitrage spreads and their correlation structure, and a restricted sample period potentially impacting the study's capture of short-term trends were acknowledged. According to (James & Menzies, 2022a), the high volatility and fragmentation inherent in the cryptocurrency market can exacerbate these inconsistencies, leading to larger and more frequent price differentials. The price variability in different cryptocurrency marketplaces presents a compelling study area, particularly to understand the factors contributing to price inconsistencies across major exchanges. (Dimpfl & Peter, 2021) investigated the role of microstructure noise and its impact on price inconsistencies across cryptocurrency exchanges. Their findings showed that cryptocurrency markets have a higher degree of microstructure noise than traditional stock exchanges. This high noise level significantly contributes to price inconsistencies.

According to (Aloosh & Ouzan, 2020), cryptocurrency markets exhibit a significant small price bias. The authors suggest that investors' reactions to news events may vary based on the current price level. (Borri & Shakhnov, 2020) investigated the occurrence of arbitrage opportunities in cryptocurrency markets, which are instances where the bid price on one exchange exceeds the ask price on another. Their findings revealed that these situations, although brief, could yield profits. Their research attributed these price discrepancies primarily to exchange-specific characteristics, such as liquidity, transaction costs, withdrawal restrictions, and geographic factors, especially the prominent role of South Korean exchanges. The study also observed a declining trend in these opportunities, indicating a growing efficiency in the cryptocurrency markets.

(Alexander & Heck, 2020) found that unregulated exchanges often diverge from standard market prices due to lower trading volumes and increased speculative trading, given the lack of required customer identification. Standard financial regulations, such as know-your-customer rules, can profoundly stabilize the cryptocurrency market by curtailing illicit activities like money laundering and fraud in less regulated exchanges and boosting investor confidence, which draws more institutional participation (Pieters & Vivanco, 2017).

(Shynkevich, 2023) investigated the price deviations and returns to arbitrage in the Bitcoin and Ethereum markets across six major cryptocurrency exchanges (BitStamp, Kraken, Binance, Bitfinex, Bittrex, and Gemini). They examined whether the prices of these cryptocurrencies on different exchanges deviate significantly from each other and whether such deviations allow for profitable trading opportunities. Results showed significant and recurring price deviations, more pronounced between countries than within, underlining a segmented market. The main reasons for price discrepancies were differences in liquidity, trading volume, cross-trading availability between cryptocurrencies, and the introduction of futures markets. However, the results indicate a limitation pertaining to Ethereum's distinctive dynamics. The price deviations for Ethereum were higher on average, more volatile, and displayed differences in the percentages of positive and negative occurrences, which affect the generalizability of the results.

Cryptocurrency exchanges, unlike traditional stock exchanges operating under strict regulatory guidelines, function as profit-oriented organizations that provide centralized exchange features using advanced technology-based platforms. These unique business models and operating conditions, inclusive of exchange-specific factors and profit dynamics, significantly influence the price formation of cryptocurrencies on their platforms (Kerr et al., 2023). Prices for Bitcoin and Altcoins can vary across different exchanges due to several factors such as liquidity, media influence, market sentiment, global events, regulatory constraints, exchange fees, and market size. While the impact of factors like Bid-Ask Spread, Twitter sentiment, Maker/Taker Fees, and Transaction Volume on price inconsistencies have been partially examined in existing studies, their cumulative effect remains underexplored (Aloosh & Ouzan, 2020; Borri & Shakhnov, 2020; Cohen, 2021; Pieters & Vivanco, 2017). The role of technical indicators is another essential aspect that has not been addressed in past research on evaluating price inconsistencies among cryptocurrency exchanges. Technical indicators are mathematical tools that are used by traders to make buy and sell decisions based on the historical price data by identifying the trends and patterns in the time series (L. Liu, 2019). There are many technical indicators that can reveal distinct trends and volatility patterns across cryptocurrency exchanges and explain the price difference between different exchanges. These technical indicators can be momentum-based, volume-based, or based on other market dynamics (Tripathi & Sharma, 2022). Variations in volumes or trading patterns across exchanges, influenced by these technical indicators, could potentially lead to diverging price trends, warranting a detailed exploration. For example, indicators such as Moving Averages or the MACD (Moving Average Convergence Divergence) enable the identification of upward, downward, or sideways market trends, which can influence trading decisions and possibly impact price disparities across exchanges, given each platform's distinct trader behavior and sentiment. Further, the assessment of market momentum, facilitated by indicators like the Relative Strength Index (RSI) or the Stochastic Oscillator, can provide insights into the overbought or oversold conditions of a cryptocurrency. Such insights can significantly affect when traders enter or exit a trade, leading to potential fluctuations in the price consistency across different exchanges. Moreover, gauging market volatility through tools such as Bollinger Bands or the Average True Range (ATR) provides insights into the risk and potential reward associated with a cryptocurrency's price. High volatility, often signifying higher risk but also higher potential reward, can lead to significant price inconsistencies across exchanges due to varying trader risk appetites. Investigating these overlooked areas is interesting and necessary to understand price inconsistencies across cryptocurrency exchanges comprehensively.

Major exchanges like Coinbase, Kraken, and Binance, each with their distinctive operational characteristics, geographical locations, regulatory compliances, and market dynamics, considerably sway the prices of cryptocurrencies on their platforms. These price variations subsequently lead to marked inconsistencies across exchanges, posing significant profitability concerns for traders and investors. Notwithstanding the relevance of this topic, research identifying the intertwined effects of these factors on price inconsistencies is scarce. Therefore, this study raises two crucial academic questions /unaccounted for in prior research. The first question is, what factors contribute to price inconsistencies across major cryptocurrency exchanges for Bitcoin and Altcoins? These factors are inclusive but not limited to various technical indicators, GitHub commits, liquidity, media influence, market sentiment, global events, regulatory constraints, exchange fees, and market size. The second question this study proposes is how does the influence of these factors vary across different cryptocurrency exchanges? This line of inquiry extends the ongoing debate on the dynamics of cryptocurrency price inconsistencies (Brauneis et al., 2022; Dimpfl & Peter, 2021; Koutmos & Wei, 2023; Pieters & Vivanco, 2017) to offer a more comprehensive understanding of this complex market phenomenon by accounting for previously overlooked factors.

2.3.1 Summary of Research Gaps for Objective 3

i. **Price Inconsistencies Across Cryptocurrency Exchanges**: The cryptocurrency market, despite its global accessibility and modern technology-driven platforms, experiences significant pricing inconsistencies across different exchanges. Existing research suggests that these inconsistencies arise from factors such as the absence of a centralized regulatory authority, varying liquidity levels, distinct market characteristics, and diverse regulatory practices. However, the literature notes limited exploration into the precise mechanisms driving cross-exchange price discrepancies. Additional investigation to thoroughly identify and examine the specific factors contributing to deviations is warranted.

ii. **Exchange-Specific Factors**: Each cryptocurrency exchange operates in a unique environment, hosting distinct user bases, security incidents, fee structures, and potential market manipulation. These exchange-specific factors can significantly impact price inconsistency, but their individual contributions have not been thoroughly examined. Developing tailored models and methods for each exchange represents a crucial research gap in the existing research.

iii. **Impact of Regional Factors**: Noted price divergences across countries imply market segmentation along regional lines. Differences in liquidity, exchange risk, and governance risk across regions likely modulate pricing but the literature notes their limited examination.

iv. **Overlooked Factors**: The cumulative impact of overlooked indicators including GitHub activity on the specific cryptocurrencies, technical factors, development activity, media influence, sentiment, global events, regulations, fees, and market breadth on pricing inconsistencies remains understudied. Comprehensively investigating their collective influence across platforms merits exploration moving forwards. Unregulated exchanges frequently demonstrate pricing divergences linked to factors such as lower volumes and heightened speculation. However, research into the potential stabilizing impact of regulations around risk and transparency remains relatively unexplored within literature. Examining this relationship merits attention.

v. **Lack of Methodological Exploration**: Previous literature notes limited diversity in terms of methodological approaches employed to analyze pricing inconsistencies thus far, including structural models, linear techniques, and microstructure examinations. Adopting more advanced and varied techniques capable of capturing intricate dynamics and spikes holds promise for developing a more nuanced understanding. Developing exchange-specific forecasting models accounting for unique characteristics including liquidity, user profiles and regional trends represents a gap within literature and an opportunity for further enhancing pricing consistency comprehension.

vi. **Impact of Institutional Participation**: Understanding the impact of participation by institutional investors coupled with increasing regulatory scrutiny offers additional, under-researched avenues for examining drivers of pricing inconsistencies through stabilizing market impacts.

2.4 Liquidity Analysis of Cryptocurrency Markets and Traditional Stock Indices

Liquidity is a fundamental concept of any financial market. It refers to the ease with which an asset can be bought or sold without affecting its original price. A market is said to be highly liquid if assets can be sold at any time with little or no price impact. Conversely, an illiquid market is one in which selling assets can lead to significant price changes due to a lack of buyers or sellers (Díaz & Escribano, 2020). In traditional finance, leading stock exchanges are often considered models of liquid markets. These markets, with their vast numbers of diverse participants, centralized price discovery mechanisms, and extensive regulations, ensure a smooth transaction process, resulting in high levels of liquidity. This liquidity allows for seamless trading experiences, enabling investors to buy and sell assets easily and quickly. Furthermore, the liquidity of a stock market reflects its stability, market efficiency and overall maturity (Hammami & Boujelbene, 2022). A high liquidity indicates a robust and well-functioning market, facilitating the accurate valuation of assets, minimizing transaction costs, and reducing price volatility (Leirvik, 2022).

Liquidity's role is not only limited to facilitating transactions. It has broader implications for market stability, fairness, and resilience. High liquidity can help absorb large trades and sudden market shocks, ensuring price stability and preventing excessive fluctuations (Yamada, 2022). Moreover, liquid markets are fairer as they prevent large players from manipulating prices to their advantage. However, the advent of cryptocurrencies and their respective exchanges has added a new dimension to our understanding of liquidity. Past research on the liquidity of cryptocurrency markets reveals a significant disparity in liquidity across cryptocurrency exchanges, creating a unique trading environment (Brauneis et al., 2018; Giudici & Pagnottoni, 2019; Park & Chai, 2020). Unlike traditional stock exchanges, cryptocurrency exchanges operate in a highly decentralized environment. The lack of a centralized authority and the high volatility inherent in Bitcoin and other cryptocurrencies contributes to the inefficiency observed in the cryptocurrency markets (Leirvik, 2022). Cryptocurrency exchanges, much like their traditional counterparts, also depend on liquidity for efficient operations (Wei, 2018). However, the factors influencing liquidity in these markets can be quite different from traditional markets due to their unique characteristics. For instance, the lack of a centralized pricing mechanism can lead to significant price disparities across different exchanges. Furthermore, the inherent volatility of cryptocurrencies can cause drastic price changes, affecting their liquidity (Pieters & Vivanco, 2017). Furthermore, the volatility of cryptocurrency exchanges, can lead to lower liquidity and higher investment risk (Al-Yahyaee et al., 2020; Hammami & Boujelbene, 2022).

In the rapidly evolving landscape of cryptocurrency trading, understanding the liquidity of cryptocurrency exchanges and benchmarking them against traditional stock markets becomes a critical concern for market participants to measure the efficiency and maturity of cryptocurrency markets. The unique aspects of cryptocurrency exchanges, such as the prevalence of algorithmic trading, the potential for market manipulation, and the constant introduction of new cryptocurrencies, present distinct challenges for maintaining and assessing liquidity of these exchanges (Brauneis et al., 2022; Koutmos & Wei, 2023; Wei, 2018). There are many other factors can significantly influence the liquidity of cryptocurrency markets. A high trading volume of exchange, representing increased trading activity and substantial market participation, is positively correlated with increased liquidity (Auer & Rottmann, 2019). Additionally, the diversity market participants, market depth, and the number of exchanges listing a specific cryptocurrency all substantially contribute to its liquidity. Moreover, the geographical location of the exchange, adherence to KYC norms, and regulatory compliance influence the trader's perception of the exchange, directly impacting trader participation and market liquidity (Feinstein & Werbach, 2020).

The ease of converting a cryptocurrency into fiat currency is vital for exchange liquidity. A reliable platform that supports this conversion, along with the availability of multiple fiat currency options, enhances liquidity by widening the range of users able to trade and exchange. Cryptocurrencies with established partnerships with payment processors or widespread acceptance as a form of payment typically exhibit enhanced liquidity owing to increased trader accessibility. Furthermore, market sentiment and investor confidence significantly impact liquidity. Positive news and progressive developments associated with a cryptocurrency exchange can attract more participants, significantly boosting the exchange's liquidity (Yue et al., 2021). Moreover, Technological factors, including the scalability of the underlying blockchain technology, play an equally significant role. The high price volatility typical of cryptocurrencies can attract or deter traders, impacting liquidity (Brauneis et al., 2021a). Additionally, trading costs, including transaction fees and bid-ask spreads, can either promote or hamper trading activity and affect liquidity. Moreover, unlike traditional markets where liquidity levels are relatively stable, these levels can vary significantly in the cryptocurrency market, adding another layer of complexity for investors. This unpredictability affects the current trading scenario and can impact future trading decisions, making it a critical aspect of market dynamics to monitor and understand. Existing literature largely focuses on price forecasting and volatility in the cryptocurrency market, leaving areas like liquidity variability and evaluation methodologies under-studied. While some studies (Brauneis et al., 2021b; Loi, 2017; Wei, 2018) have examined the liquidity characteristics of cryptocurrency exchanges, a specific comparative assessment of liquidity with that of leading stock indices remains largely underexplored to the best of our knowledge.

(Corbet et al., 2019) compared Bitcoin's liquidity to that of gold and the U.S. dollar and discovered that Bitcoin's liquidity is inferior to that of these traditional assets. Conversely, some works presented a contrasting finding, suggesting an improving liquidity condition in the cryptocurrency market. (Sensoy, 2019) showed an upward trend in Bitcoin's liquidity over time, positing that its liquidity might eventually surpass traditional assets given the right conditions and regulatory support. Increasing trading activity and user participation in an exchange improves liquidity, enhancing trade execution at desired prices. Nevertheless, liquidity fluctuates with market conditions and sentiment toward cryptocurrencies. A fair assessment necessitates analyzing liquidity over various periods and conditions rather than relying on a single snapshot, which may not reflect true liquidity. However, most works only consider a single snapshot of liquidity, which may not accurately reflect the true liquidity of an exchange (Cortez et al., 2021; Doumenis et al., 2021; Loi, 2017; Takaishi & Adachi, 2020). In addition, while the existing literature provides valuable insights into Bitcoin's liquidity dynamics, many works fall short of including the broad range of cryptocurrencies into the analysis (Brauneis et al., 2022; Loi, 2017; Takaishi & Adachi, 2020). Cryptocurrency assets, besides Bitcoin, such as Ethereum, Ripple, Litecoin, and many others, contribute significantly to the overall crypto market capitalization and display distinct liquidity characteristics that warrant separate investigation. Furthermore, most existing studies have primarily focused on single-market liquidity (Corbet et al., 2019; Cortez et al., 2021; Doumenis et al., 2021; Loi, 2017; Takaishi & Adachi, 2020), leaving cross-market liquidity unexplored to a significant extent. Cross-market liquidity, here refers to comparing the liquidity of cryptocurrency markets with traditional stock markets.

As cryptocurrencies gain broader acceptance into the mainstream finance, studying the liquidity of cryptocurrency marketplaces transitions from an academic exercise to a practical necessity. In this work, we examine the liquidity of leading cryptocurrency exchanges such as Kraken, Coinbase, and Binance, comparing these measurements against leading stock indices to assess the efficiency and maturity of cryptocurrency markets. Our research endeavors to interpret liquidity from the standpoint of investors and traders, who require this knowledge for informed decision-making regarding trading venue selection and risk management strategies. We aim to evaluate the liquidity using widely accepted liquidity measurement tools like the Martin Liquidity Index and Amihud's Illiquidity ratio for capturing the depth and immediacy of market liquidity in the crypto context need to be critically analyzed (Coen & de La Bruslerie, 2019; Darolles et al., 2011; Hammami & Boujelbene, 2022). From a practical implication standpoint, this research aids market participants for a prospective future where integration between traditional and cryptocurrency markets is tighter, prompting shared liquidity considerations. For regulators and policy-makers, a deep understanding of liquidity in cryptocurrency exchanges can guide the formulation of effective regulatory frameworks. As regulators across the globe grapple with the challenge of integrating cryptocurrencies into their financial systems, liquidity considerations should form a crucial part of their approach. Effective regulations can help mitigate liquidity risk, prevent market manipulation, and protect investors. Moreover, the study of liquidity in cryptocurrency exchanges can provide insights into the broader financial ecosystem's evolution. The intersection of traditional finance and cryptocurrencies presents an unprecedented opportunity to explore new models of market functioning, financial intermediation, and management of systemic risk. Understanding liquidity in this context can lead to the development of more resilient, efficient, and inclusive financial systems. We aim to address the research gaps in the literature by comparing the liquidity of major cryptocurrency exchanges with that of traditional stock exchanges. Utilizing different liquidity measures we focus on the top three volume traded cryptocurrency exchanges based and compare them with leading traditional markets such as NYSE, NASDAQ, and BSE SENSEX. The primary objectives of the study are threefold. Firstly, it seeks to analyze how liquidity in major cryptocurrency exchanges compares with that of traditional stock exchanges. Secondly, it aims to identify which of the major cryptocurrency exchanges exhibit the highest liquidity. Finally, we aim to assess liquidity changes over time to investigate how the cross-exchange liquidity has evolved. By doing so, the study seeks to provide investors and traders with vital information that can guide their decision-making process regarding where to trade.

2.4 Summary of Research Gaps for Objective 4

i. **Cross-Market Comparative Liquidity Analysis**: The existing literature lacks a detailed comparison of liquidity between major cryptocurrency exchanges and traditional stock indices. While differences in liquidity characteristics are acknowledged, a systematic study is necessary to understand how liquidity in cryptocurrency exchanges diverges from that in traditional markets. The comparison of liquidity between cryptocurrency markets and traditional stock markets (e.g., NYSE, NASDAQ, BSE SENSEX) is not extensively explored. Analyzing how liquidity in cryptocurrency exchanges compares with that in established stock exchanges is crucial for evaluating the maturity and efficiency of cryptocurrency markets.

ii. **Focus predominantly on Bitcoin**: Most existing studies focus primarily on Bitcoin, neglecting other cryptocurrencies like Ethereum, Ripple, and Litecoin. These assets contribute significantly to the market and may show different liquidity characteristics. A broader analysis including a variety of cryptocurrencies is essential for a complete understanding of the crypto market's liquidity.

iii. **Gap in Liquidity Measurement Tools**: The existing literature review reveals a gap in the comprehensive application of liquidity measurement tools in the context of cryptocurrency exchanges. This thesis aims to bridge this gap by introducing and evaluating additional methods such as the Volume Weighted Average Price (VWAP), Bid-Ask Spread, Turnover Ratio, Kyle's Lambda, and the Effective Spread, and comparing them against the baseline Martin Liquidity Index (MLI) measure. The rationale behind this comparative analysis is to identify the most effective tools for assessing liquidity in the volatile and diverse environment of cryptocurrency markets. This approach will enhance our understanding of liquidity dynamics in cryptocurrency exchanges and provide a more robust framework for liquidity assessment in this evolving financial landscape.

iv. **Longitudinal Analysis**: Liquidity changes over time, yet most studies use single snapshots, potentially missing the dynamic nature of liquidity. A longitudinal approach is necessary to understand how liquidity evolves under varying market conditions.

v. **Regulatory Considerations and Financial Ecosystem Evolution**: As traditional and cryptocurrency markets become more integrated, understanding liquidity is vital for developing effective regulatory frameworks. Research should explore how these frameworks can mitigate liquidity risks, prevent market manipulation, and protect investors. The intersection of traditional finance and cryptocurrencies offers opportunities for new market functioning models and financial intermediation. Studying liquidity in this context can aid in developing more resilient and inclusive financial systems.

Chapter 3

Data and Research Methodologies

This chapter provides a detailed description of the methodologies and datasets employed in this research. The initial section outlines the datasets chosen based on the distinct requirements of each research objective. The chapter then progresses to discuss the research methods. For each objective, we present schematic diagrams to explain the overall methodological approach. In addition, we describe the specific techniques employed, the specifics of data preprocessing, the algorithms, and the rationale behind the methods chosen.

3.1 Data

The choice of data directly influences the effectiveness of models and methodologies applied. Recognizing that each objective poses its unique set of research questions, the data needs to be both specific and prepared appropriately through tailored preprocessing. This section presents the datasets carefully chosen to cater to the specificities of each research objective and its inherent research queries.

3.1.1 Dataset for Objectives 1 and 2

To fulfill the first two objectives of this work, which involve analyzing and forecasting the exchange prices of Bitcoin and Altcoins and determining the price formation factors, we require an exhaustive dataset for each cryptocurrency that reflects the multifaceted nature of these cryptocurrencies. Given the unique characteristics of each asset, it is imperative to consider individual datasets for Bitcoin and each Altcoin. While the foundational principles of these cryptocurrencies are derived from Bitcoin, each Altcoins' Blockchain has evolved distinctively, incorporating modifications and enhancements that may influence their price dynamics differently. Therefore, a careful analysis of these price trends necessitates the analysis of individual datasets for Bitcoin and each Altcoin.

The evaluation of fundamental and technical indicators is crucial for analyzing and forecasting the exchange prices of Bitcoin and Altcoins and determining their price formation factors. Fundamental indicators provide insights into the intrinsic value of a cryptocurrency. They encompass factors such as the transaction fees, hash rate, mining difficulty, transactions per day, macro-economic factors, social media trends and many such variables. These indicators reflect a cryptocurrency's underlying health and utility, critical determinants of its long-term value. By analyzing these indicators, we can better understand the fundamental drivers of cryptocurrency prices.

On the other hand, technical indicators focus on price movements and trading volumes. They are primarily used to identify and predict short-term price trends based on historical trading data. Technical indicators can provide valuable insights into market sentiment and potential price reversals, which is particularly useful in the highly volatile cryptocurrency market.

We investigate cryptocurrency market prices by first analyzing fundamental indicators and subsequently, technical indicators. After these individual assessments, we combine the insights from both indicators to gain a more comprehensive understanding. This three-layered approach allows us to understand both the long-term value and short-term price movements of Bitcoin and Altcoins, thereby enhancing the accuracy of our price forecasts and the robustness of our analysis on price formation factors.

We use daily Close price of Bitcoin and Altcoins as the dependent variable. Table 1 summarizes the list of all independent variables. Macroeconomic variables, classified under Type ID 1, are downloaded from http://www.investing.com. Public sentiments and attention towards the respective coins are reflected by Google Trends data and Twitter data are categorized under Type ID 2 and sourced from https://trends.google.co.in/trends and https://bitinfocharts.com respectively. Cryptocurrency prices, categorized under Type ID 3, are accessed from http://www.coinmarketcap.com through its API for Bitcoin and Altcoins (Ethereum, Ripple, Dash, and Litecoin). Technical indicators, spanning Type IDs 4-7, are derived from Open, Low, High, Close, and Volume data using the Technical Analysis Library Python package. All calculations formulae for these technical indicators are documented in detail Technical Analysis Library documentation at https://technical-analysis-library-in-python.readthedocs.io/en/latest/ta.html. A comprehensive data dictionary, providing a detailed description of each variable, is available at the following links in excel, json and xml formats.

Table 2.1 Summary of the Fundamental and Technical indicators used as independent variables for Objective 1 and Objective 2. All data variables are composed of daily observations

Type ID	**Indicator Type**	**Category**	**Variable Name**
1	Fundamental	Macroeconomic	CNY_USD_data_Close, EUR_USD_data_Close, GBP_USD_data_Close, RUB_USD_data_Close, WTI_crudeoil_data_Close, SP500_Close, NASDAQ_Close, Gold_Close, Dow_Jones30_Close, US_10Y_Treasury_rate_Close, US_Initial_Claims_Close
2	Fundamental	Speculative	google_trends, Tweets
3	Technical	Price	Open, High, Low, Close, Volume
4	Technical	Volume	volume_adi, volume_obv, volume_cmf, volume_fi, volume_em, volume_sma_em, volume_vpt, volume_vwap, volume_mfi, volume_nvi
5	Technical	Volatility	volatility_bbm, volatility_bbh, volatility_bbl, volatility_bbw, volatility_bbp, volatility_bbhi, volatility_bbli, volatility_kcc, volatility_kch, volatility_kcl, volatility_kcw, volatility_kcp, volatility_kchi, volatility_kcli, volatility_dcl, volatility_dch, volatility_dcm, volatility_dcw, volatility_dcp, volatility_atr, volatility_ui
6	Technical	Trend	trend_macd, trend_macd_signal, trend_macd_diff, trend_sma_fast, trend_sma_slow, trend_ema_fast, trend_ema_slow, trend_vortex_ind_pos, trend_vortex_ind_neg, trend_vortex_ind_diff, trend_trix, trend_mass_index, trend_dpo, trend_kst, trend_kst_sig, trend_kst_diff, trend_ichimoku_conv, trend_ichimoku_base, trend_ichimoku_a, trend_ichimoku_b, trend_stc, trend_adx, trend_adx_pos, trend_adx_neg, trend_cci, trend_visual_ichimoku_a, trend_visual_ichimoku_b, trend_aroon_up, trend_aroon_down, trend_aroon_ind, trend_psar_up,

			trend_psar_down, trend_psar_up_indicator, trend_psar_down_indicator
7	Technical	Momentum	momentum_rsi, momentum_stoch_rsi, momentum_stoch_rsi_k, momentum_stoch_rsi_d, momentum_tsi, momentum_uo, momentum_stoch, momentum_stoch_signal, momentum_wr, momentum_ao, momentum_roc, momentum_ppo, momentum_ppo_signal, momentum_ppo_hist, momentum_pvo, momentum_pvo_signal, momentum_pvo_hist, momentum_kama

Table 2.1 provides a detailed overview of blockchain variables for Bitcoin and four Altcoins: Ethereum, Ripple, Litecoin, and DASH. These variables serve as independent variables for Objectives 1 and 2 of our research. The historical data for all Altcoins' Blockchain variables are obtained from https://bitinfocharts.com, while the historical data for Bitcoin's Blockchain is obtained from http://www.blockchain.com. For all Altcoins, we also consider the Opening, Low, High, and Close prices of Bitcoin (Type ID 3) as the change in these prices could influence the Close prices of Altcoins. The Ethereum data is sourced from a publicly accessible dataset provided by AWS (https://aws.amazon.com/blogs/database/access-bitcoin-and-ethereum-open-datasets-for-cross-chain-analytics).

Although Table 2.1 and 2.2 describe specific variables, each corresponds to a separate dataset for a distinct cryptocurrency in our analytical process. By examining these tables, we can analyze these variables and utilize them in forecasting models, allowing us to examine the behavior and characteristics of the five cryptocurrencies under study.

Table 2.2 A descriptive overview of Blockchain variables for Bitcoin and Altcoins (Ethereum, Ripple, Litecoin and DASH) used as independent variables for Objective 1 and Objective 2.

Type ID	Category	Coin	Feature Name	Description
1		Bitcoin	Market Value to Realized Value	Ratio of market value to realized value of Bitcoin
2		Bitcoin	Network Value to Transactions	Ratio of network value to transactions of Bitcoin
3		Bitcoin, Altcoins	Average Price per Day	The average trading price of the cryptocurrency on the blockchain in a day.
4	**Market Signals**	Bitcoin, Altcoins	Market Capitalization	The total market value of the cryptocurrency on the Blockchain.
5	**Mining Information**	Bitcoin	Cost of Transaction	Cost of processing Bitcoin transactions
6		Bitcoin	Cost as a Percentage of Transaction Volume	Percentage of transaction volume that goes towards processing costs
7		Bitcoin	Fees per Transaction (in USD)	Fees per Bitcoin transaction in USD
8		Bitcoin	Miners' Revenue (in USD)	Revenue of Bitcoin miners in USD
9		Bitcoin	Network Difficulty	Difficulty of mining Bitcoin
10		Bitcoin	Total Hash Rate (in TH/s)	Total hash rate of Bitcoin network in tera hashes per second
11		Bitcoin	Total Transaction Fees (in BTC)	Total transaction fees in Bitcoin
12		Bitcoin	Total Transaction Fees (in USD)	Total transaction fees in USD
13		Bitcoin, Altcoins	Average Mining Difficulty per Day	The average difficulty level for mining new blocks on the blockchain in a day.
14		Bitcoin, Altcoins	Average Hash rate per Day	The average computational power used in mining and processing transactions on the blockchain in a day.
15		Bitcoin, Altcoins	Mining Profitability	The profitability of mining new blocks on the blockchain, considering costs and rewards.
16	**Mining Information**	Bitcoin, Altcoins	Average Fee as a Percentage of Total Block Reward	The average percentage of transaction fees in the total block reward on the blockchain in a day.
17		Bitcoin	Confirmed Transactions per Day	Number of confirmed Bitcoin transactions per day
18		Bitcoin	Estimated Transaction Value (in BTC)	Estimated value of Bitcoin transactions
19	**Network**	Bitcoin	Estimated Transaction Value (in USD)	Estimated value of Bitcoin transactions in USD
20	**Activity**	Bitcoin	Output Value per Day	Output value of Bitcoin per day

21		Bitcoin	Transactions Excluding Popular Addresses	Number of Bitcoin transactions excluding those involving popular addresses
22		Bitcoin	Unique Addresses Used	Number of unique Bitcoin addresses used
23		Bitcoin	Unspent Transaction Outputs	Number of unspent Bitcoin transaction outputs
24		Bitcoin, Altcoins	Daily Transactions Count	The total number of transactions conducted on the blockchain in a day.
25		Bitcoin, Altcoins	Average Block Size	The average size of blocks mined on the blockchain in a day.
26		Bitcoin, Altcoins	Unique Addresses Used per Day	The number of unique addresses involved in transactions on the blockchain in a day.
27	**Wallet Information**	Bitcoin	Blockchain.com Wallets Activity	Activity related to blockchain.com wallets
28		Bitcoin, Altcoins	Sent Coins per Day	The total amount of cryptocurrency sent in transactions on the blockchain in a day.
29		Bitcoin, Altcoins	Average Transaction Fees per Day	The average transaction fees on the blockchain in a day.
30		Bitcoin, Altcoins	Average Block Confirmation Time	The average time taken to confirm a new block on the blockchain in a day.
31	**Transaction Details**	Bitcoin, Altcoins	Average Transaction Value (in USD) per Day	The average value of transactions on the blockchain in USD in a day.
32	**Wealth Distribution**	Bitcoin, Altcoins	Top 100 Richest Addresses' Holdings	The total wealth held by the top 100 richest addresses on the blockchain.
33		Altcoins	Hash count (ETH)	The number of hashes computed per second by miners in the Ethereum network.
34	**Ethereum Specific Data**	Altcoins	Token transfers (ETH)	The transfers of Ethereum-based tokens between addresses.
35			Block count (ETH)	The number of blocks that have been added to the Ethereum blockchain.

The first interval encompasses the early development and evolution of the cryptocurrency market, including its response to significant events such as the Mt. Gox incident in 2014. It is important to highlight that not all cryptocurrencies have data extending back to the earliest phases of the market. Therefore, the analysis of certain altcoins commences with their respective launch dates on www.coinmarketcap.com. For the phase encompassed by Interval 2, which spans from the respective inception dates of each cryptocurrency until April 30, 2017, the market underwent significant transformations. This phase traces the market's trajectory post the Mt. Gox crash, marked by increased regulatory scrutiny, the establishment of more mature market dynamics, and a steady recovery of prices. It includes a crucial time period of resilience and adjustment as cryptocurrencies worked towards gaining stability, user adoption and confronting evolving external perceptions and intrinsic challenges. The cryptocurrency market underwent significant changes during the third interval, from 2017 to 2021. Bitcoin and several other cryptocurrencies reached record highs in December 2017 and November 2021. These peaks did not just mark this period but also saw notable price corrections. The fourth interval examines the period after 2017, characterized by several price surges and corrections periods. This stage is especially significant as cryptocurrencies began to receive wider recognition and attention from mainstream financial institutions during this phase. Specific data intervals for each cryptocurrency, detailing their start and end dates, can be referenced in Table 3, 4, 5, 6 and 7. These tables provide a structured breakdown for Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), Ripple (XRP), and Dash.

Table 3. Data intervals considered for Bitcoin (BTC)

Intervals	Start Date	End Date
Interval 1	29-Apr-13	31-Jul-16
Interval 2	29-Apr-13	30-Apr-17
Interval 3	29-Apr-13	31-Dec-17
Interval 4	02-Jan-17	02-Apr-23

Table 4. Data intervals considered for Ethereum (ETH)

Intervals	Start Date	End Date
Interval 1	7-Aug-15	31-Jul-16
Interval 2	7-Aug-15	30-Apr-17
Interval 3	7-Aug-15	31-Dec-17
Interval 4	02-Jan-17	02-Apr-23

Table 5. Data intervals considered for Litecoin (LTC)

Intervals	Start Date	End Date
Interval 1	29-Apr-13	31-Jul-16
Interval 2	29-Apr-13	30-Apr-17
Interval 3	29-Apr-13	31-Dec-17
Interval 4	02-Jan-17	02-Apr-23

Table 6. Data intervals considered for Ripple (XRP)

Intervals	Start Date	End Date
Interval 1	04-Aug-13	31-Jul-16
Interval 2	04-Aug-13	30-Apr-17
Interval 3	04-Aug-13	31-Dec-17
Interval 4	02-Jan-17	02-Apr-23

Table 7. Data intervals considered for Dash

Intervals	Start Date	End Date
Interval 1	14-Feb-14	31-Jul-16
Interval 2	14-Feb-14	30-Apr-17
Interval 3	14-Feb-14	31-Dec-17
Interval 4	2-Jan-17	2-Apr-23

3.1.2 Dataset for Objective 3

The third and fourth objectives of this study involve identifying the factors responsible for price inconsistencies across different major cryptocurrency exchanges and comparing the liquidity of cryptocurrency exchanges with the liquidities of different sizes of selected stocks using the Martin Liquidity Index (MLI). These objectives necessitate the collection of additional datasets that capture the trading dynamics on various cryptocurrency exchanges and the liquidity characteristics of selected stocks.

We collected daily data from three prominent cryptocurrency exchanges: Binance, Coinbase, and Kraken from 01-Jan-17 to 10-Apr-23 for our analysis. For consistency with previous objectives, we defined the start of a trading day as the first transaction after 00:00:00 Coordinated Universal Time (UTC) and the end as the last transaction before 23:59:59 UTC for our analysis. The variables used to create the dataset for each cryptocurrency on each of the three exchanges are listed in Table 2.2. In addition to the variables presented in Table 2.2, we also include in our analysis the technical indicators defined in Type IDs 4 through 7 from Table 2.1 as additional independent variables.

Table 8. Independent variables utilized for evaluating discrepancies in cryptocurrency prices across major exchanges

S.No.	Variable Name	Description
1	open	Opening price of the cryptocurrency
2	high	Highest price of the cryptocurrency in each period
3	low	Lowest price of the cryptocurrency in each period
4	close	Closing price of the cryptocurrency
5	volume	Total trading volume of the cryptocurrency in each period
6	github_commits	Number of commits made to the cryptocurrency's GitHub repository
7	google_trends	Google search interest for the cryptocurrency
8	tweets	Number of tweets mentioning the cryptocurrency
9	Maker Fee	Fee charged to the maker in a transaction
10	Transaction Fee	Fee charged for conducting a transaction
11	Bid-Ask spread	Difference between Bid and Ask Price
11	AML_KYC_required	Indicator if Anti-Money Laundering and Know Your Customer regulations are required

For the third objective, the independent variable of interest is the price inconsistency across different cryptocurrency exchanges. This variable is computed as follows:

Let $P_{d,c,e}$ represent the closing price for a particular date d, coin c, and exchange e. The average closing price $P_{avg_{d,c}}$ for date d and coin c across all exchanges is calculated as:

$$P_{avg_{d,c}} = \frac{1}{N}\sum_{e} P_{d,c,e} \tag{1}$$

where the sum is over all exchanges e, and N is the total number of exchanges.

The price inconsistency $I_{d,c,e}$ for a particular date d, coin c, and exchange e is then calculated as the absolute difference between the closing price and the average closing price:

$$I_{d,c,e} = \left|P_{d,c,e} - P_{avg_{d,c}}\right| \tag{2}$$

This results in a measure of price inconsistency for each coin on each exchange for each date. The symbols $P_{avg_{d,c}}$ and $I_{d,c,e}$ represent the average price and price inconsistency respectively.

3.1.3 Dataset for Objective 4

The liquidity of cryptocurrency exchanges is derived directly from the currency pairings with the highest trading volume. We use indices as proxies for the overall liquidity of traditional stock exchanges, with data sourced from the Yahoo Finance API. Specifically, we utilize the NYSE Composite Index, which represents all common stocks on the NYSE, includes more than 2,000 stocks from various industries, and serves as a vital benchmark for the US stock market. The NASDAQ Composite, which consists of over 3,000 companies, primarily represents the technology and growth sectors. In addition, we use the BSE SENSEX, which comprises 30 of the largest corporations listed on the Bombay Stock Exchange and serves as the primary benchmark for the Indian equity market.

Our cryptocurrency analysis focused on the highest volume-traded currency pairs from each exchange. On Kraken, we considered the top 150 currency pairs, representing 98.48% of the total trading volume. For Coinbase, the analysis was based on the top 167 currency pairs, accounting for 99.02% of its total volume. Binance's data was drawn from the top 200 currency pairs, making up 94.63% of its trading volume. Notably, the primary data for Binance was sourced from binance.com, the primary global platform, excluding region-specific versions such as binance.co.uk and binance.us. This approach ensures that our analysis is anchored on stable assets, minimizing the influence of volatile or speculative coins. We obtain daily transaction records of price values with Open, High, Low, Close, Volume, and Number of Trades (OHLCVT) data for investigating the cryptocurrency exchange data. We defined the start of a trading day as the first transaction after 00:00:00 Coordinated Universal Time (UTC) and the end as 23:59:59 UTC for all cryptocurrency-related data given that Bitcoin exchanges operate continuously, seven days a week, with no typical daily opening and closing times.

We use the intervals listed in Table 9 for both traditional stock market data and cryptocurrency market data. Interval 1, spanning from 08-Sep-19 to 01-Jan-20, captures the late stages of the early adoption phase and the onset of mainstream adoption for cryptocurrencies, marked by Bitcoin's significant bull run and the ICO boom. In contrast, traditional markets during this period experienced robust global economic growth, with stock markets reaching all-time highs. Interval 2, from 01-Jan-20 to 01-Apr-23, encompasses the rise of DeFi platforms in the cryptocurrency market and increased institutional interest, set against the backdrop of the COVID-19 pandemic's onset and its profound impact on traditional markets. Central banks' monetary easing and governments' fiscal stimulus packages marked this period, leading to a strong market recovery. Interval 3, extending from 08-Sep-19 to 01-Apr-23, offers a comprehensive view of both markets, capturing the entire journey of the crypto market from its early adoption phase to its more matured state, while also reflecting the traditional markets' response to various economic challenges and recoveries. The data on the Binance API was only available as of 8-Sep-2019, although we acquired data from the Kraken and Coinbase Pro exchanges as far back as 1-Jan-2017. As a result, we chose the start date of September 8, 2019, for Interval 1 to allow for an accurate comparison of liquidity across all exchanges.

Table 9. Data extraction intervals for liquidity comparison across exchanges.

Interval Name	Start Date	End Date
Interval 1	08-Sep-19	01-Jan-20
Interval 2	01-Jan-20	01-Apr-23
Interval 3	08-Sep-19	01-Apr-23

For data retrieval, we used the Coinbase Pro API (https://api.pro.coinbase.com) for Coinbase, the Python krakenex library (https://pypi.org/project/krakenex) for downloading data from

Kraken, and the python-binance library (https://python-binance.readthedocs.io) for Binance. All currency pairs for each exchange for which the data was downloaded are listed here.

3.2 Research Methods

3.2.1 Objective 1

O1: To analyze the trends of exchange prices of Bitcoin and Altcoins and to forecast future exchange prices of Bitcoin and Altcoins

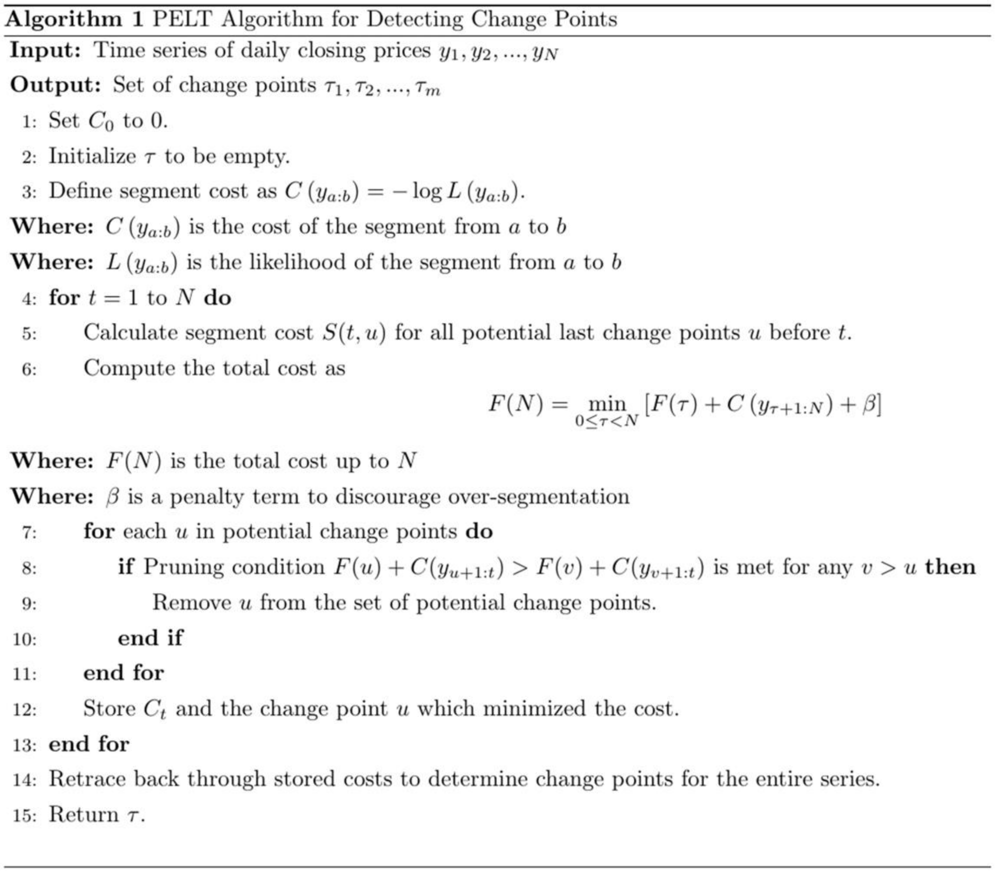
3.2.1.1 Analytical Methods for Investigating Cryptocurrency Price Patterns

In our methodology for analyzing price trends in the exchange prices of cryptocurrencies, we employ two primary techniques: Change Point Analysis using the Pruned Exact Linear Time (PELT) algorithm and Multi Fractal Detrended Fluctuation Analysis (MF-DFA). Using the PELT algorithm, we apply the Change Point Analysis to daily price data (Truong et al., 2020). It aids in identifying significant shifts in mean price levels, helping us to pinpoint and visualize alterations in price trajectories and understand market behaviors. Moreover, we analyze the descriptive statistics for the Change Points derived from the PELT algorithm for all cryptocurrencies and assess each change point in depth. Concurrently, MF-DFA is carried out on daily, weekly, and monthly price data across each predefined interval, giving a thorough perspective on price fluctuations across varying time scales.

i) Change Point Analysis Pruned Exact Linear Time (PELT) algorithm

Change point detection identifies the moments in which the probability distribution of a time series transforms. This method is especially useful for detecting anomalies and pattern variations in time series data. Change point analysis in the context of cryptocurrencies aims to estimate where the statistical properties of price data sequences change and aid in identifying shifts in price volatility, identifying intervals of increased or decreased market fluctuations, and thus providing clarity on periods marked by stability or variability. Tracking these changes across the four intervals assists in identifying the market's evolving trajectory, with fewer change points in subsequent intervals indicating a trend toward market stability. This analysis makes it possible to distinguish between long-term price trends and short-term price fluctuations, providing investors and traders with a distinct demarcation. The detected change points can also be correlated with significant external events, such as regulatory changes, technological developments, breaches of security, or broad economic shifts, enabling us to understand how such events impact cryptocurrency prices.

The Pruned Exact Linear Time (PELT) algorithm is a widely recognized algorithm for change point detection. Comparative studies have demonstrated that PELT exhibits superior performance compared to alternative change point analysis methods, like binary segmentation, and offers more expedient results than other exact search techniques (Dorcas Wambui, 2015). The Pruned Exact Linear Time (PELT) Algorithm aims to detect change points in time series data. In this specific context, we examine the daily closing prices of five major cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Dash (DASH), and Litecoin (LTC). Algorithm 1 describes the PELT algorithm for detecting change points in the input time series of each coin.



```
Algorithm 1 PELT Algorithm for Detecting Change Points
Input: Time series of daily closing prices $y_1, y_2, ..., y_N$
Output: Set of change points $\tau_1, \tau_2, ..., \tau_m$
Set $C_0$ to 0.
Initialize $\tau$ to be empty.
Define segment cost as $C(y_{a:b}) = -\log L(y_{a:b})$.
Where: $C(y_{a:b})$ is the cost of the segment from $a$ to $b$
Where: $L(y_{a:b})$ is the likelihood of the segment from $a$ to $b$
for $t = 1$ to $N$ do
    Calculate segment cost $S(t, u)$ for all potential last change points $u$ before $t$.
    Compute the total cost as
        $F(N) = \min_{0 \leq \tau < N} [F(\tau) + C(y_{\tau+1:N}) + \beta]$
Where: $F(N)$ is the total cost up to $N$
Where: $\beta$ is a penalty term to discourage over-segmentation
    for each $u$ in potential change points do
        if Pruning condition $F(u) + C(y_{u+1:t}) > F(v) + C(y_{v+1:t})$ is met for any $v > u$ then
            Remove $u$ from the set of potential change points.
        end if
    end for
    Store $C_t$ and the change point $u$ which minimized the cost.
end for
Retrace back through stored costs to determine change points for the entire series.
Return $\tau$.
```

The algorithm processes these time series, represented as $y_1, y_2, \dots, y_N$, and aims to produce a set of change points $\tau_1, \tau_2, \dots, \tau_m$. The algorithm begins by initializing the cost, C_0 to zero and formulating an empty set, τ, designated for the change points. The cost of a segment from a to b in the time series, denoted as $C(y_{a:b})$, is defined as the negative logarithm of the likelihood of that segment, represented as $L(y_{a:b})$. In the next stage, we iterate through the time series from $t = 1$ *to* t=*N*. For each instance of t, it computes the segment cost, $S(t, u)$ for all conceivable last change points, u, preceding t. It then calculates the aggregate cost up to point N using the equation:

$$F(N) = \min_{0 \leq \tau < N} [F(\tau) + C(y_{\tau+1:N}) + \beta] \tag{3}$$

Here, the term β functions as a penalty parameter. It is integrated into the algorithm to prevent excessive segmentation of the data, ensuring that only significant change points are identified. Upon ascertaining this total cost, the algorithm evaluates each proposed change point, u. If any of these points satisfy a particular pruning criterion, specifically, if the condition $F(u) + C(y_{u+1:t}) > F(v) + C(y_{v+1:t})$

holds for any v greater than u, the algorithm removes u from the potential change points set.

After this evaluation, the algorithm stores the cost, C_t, and the specific change point, u, which minimized this cost. This indicates the last change point if the final segment concludes at t. Once the entire time series has been processed, the algorithm retraces its steps using the stored costs. This action facilitates the identification of change points spanning the series. Ultimately, it yields the set τ, encapsulating these change points.

Analyzing the daily closing prices of Bitcoin, Ethereum, Ripple, Dash, and Litecoin with the PELT algorithm helps to identify key inflection points related to market changes influenced by speculation, technology, macroeconomic factors, and other events.

ii) Multi Fractal Detrended Fluctuation Analysis

Cryptocurrency price movements, reflective of their underlying market mechanics and external influences, embody several nuanced structural characteristics that traditional statistical analyses may overlook. These include temporal dependencies, where price alterations on a given day might influence or correlate with subsequent price changes. There are also "bursty behaviors", characterized by intervals of rapid price adjustments followed by periods of relative stability, indicating the market's non-uniform activity (James & Menzies, 2022b). Furthermore, cryptocurrencies display scaling behaviors, where patterns repeat themselves across varying time scales; similar behaviors might be evident in daily, weekly, and monthly price charts. Cyclic patterns are another key characteristic, with recurring upward and downward price shifts potentially tied to specific events or timings. Additionally, the asymmetries in cryptocurrency price movements should be noted, where upward price trajectories might sharply contrast with more gradual descents. Multifractal Detrended Fluctuation Analysis (MF-DFA) is a method to assess non-stationary time series, such as cryptocurrency price data (Miloş et al., 2020). MF-DFA extends the principles of Detrended Fluctuation Analysis (DFA) by enabling the identification of multifractal characteristics within the time series.

In cryptocurrency data, fractals highlight patterns that repeat at various scales. Recognizable patterns in the price dynamics of cryptocurrencies, such as Bitcoin, emerge in timeframes from hourly to monthly charts. This consistency reveals the fractal patterns inherent in cryptocurrency price data. In multifractal systems, self-similarity varies, allowing multiple fractal structures to exist where each has distinct statistical behaviors (Takaishi, 2018). Pattern variations across different parts of a time series define 'multifractality.' Given the volatile nature of cryptocurrency prices, understanding multifractality is essential for investigating the statistical properties of these Blockchain-based assets.

A key component in multifractal exploration is the Generalized Hurst Exponent, *h(q)*. Traditionally used to determine the long-term memory or autocorrelation in a time series, the Hurst exponent provides specific insights. A value around 0.5 implies a random walk, values exceeding 0.5 indicate persistent behavior, while those below 0.5 suggest anti-persistent behavior. Extended into the generalized domain, *h(q)* can be evaluated over a range of *'q'* values, probing the time series' behavior across different scales or fluctuation magnitudes. For example, negative *'q'* values emphasize large fluctuations, while positive '*q*' values highlight smaller ones. By employing MF-DFA, variations and structures within the price time series of Bitcoin and Altcoins can be systematically evaluated.

To examine the trends in Bitcoin and Altcoin time series data and conduct multifractal analysis, we structure our analysis into the following three key stages (Miloş et al., 2020):

1. In Stage 1, we begin by calculating simple returns for each period using the difference between successive prices relative to the previous period's price. In addition, we compute the log returns, which yield the natural log of the ratio between two consecutive prices.
2. In Stage 2, we conduct an STL Decomposition, segmenting our returns into three distinct components: a dominant trend, periodic fluctuations, and a residual component. This method aids in separating consistent patterns and overall trends from random fluctuations in our cryptocurrency data.
3. Finally, in Stage 3, we use multifractal detrended fluctuation analysis (MF-DFA). This involves creating a profile based on the residual data from the previous stage. We then split this profile into segments and try to fit a trend to each. We calculate the variances for each segment from these trends, enabling us to determine a critical value known as the generalized Hurst exponent. This value helps in understanding long-term patterns and behaviors in cryptocurrency data.

Stage 1: Calculating Cryptocurrency Returns

Let P_t represent the closing price of the cryptocurrency at time t. The uncompounded return for the cryptocurrency over time interval t, denoted as r_t , is given by:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{4}$$

where $P_{t-1}$is the price at time $t-1$.

The logarithmic return of the cryptocurrency for the time period t, denoted as rL_t is expressed as:

$$rL_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{5}$$

Stage 2: STL Decomposition

Based on the Wold representation theorem, we can break down the cryptocurrency returns into deterministic and stochastic parts. We use the Seasonal-Trend decomposition (STL) method with LOESS (Locally Estimated Scatterplot Smoothing) for this decomposition. The relation can be expressed as:

$$r_t = T_i + S_i + R_i \tag{6}$$

where r_t is the return at time t, T_i is the deterministic trend component, S_i is the seasonal component, and R_i is the stochastic remainder component.

Stage 3: MF-DFA

The multifractal detrended fluctuation analysis (MF-DFA) offers a robust approach to identify multifractal characteristics in time series data. In MF-DFA, we quantify the mean volatility over specific intervals and use this statistical point to compute the corresponding volatility functions.

This method determines the generalized Hurst exponents based on the volatility's power law features. A key strength of MF-DFA is its ability to identify consistent correlations within non-stationary data series.

Stage 3.1 Profile Construction

In the first stage, we formulate the profile. Let R(i) represent the stochastic remainder component from the STL decomposition, which may exhibit non-stationarity. The profile, *Y*(*j*), is constructed as given below. First, we construct the profile, $Y(j)$:

$$Y(j) = \sum_{i=1}^{j} (R(i) - \bar{R}) \tag{7}$$

where $R(i)$ for $i = 1, \ldots, N$

Stage 3.2 Segment Division

For a detailed investigation, the profile $Y(j)$ is divided into Ns segments, each of uniform size. This is formulated as:

$$Ns = \text{int}\left(\frac{N}{s}\right) \tag{8}$$

where s is the length of each segment, and *N* is the total length of the time series.

Notably, as Ns might not always be an integer multiple of s, a segment of the profile might remain unutilized. To address this, we also initiate the division from the opposite end, resulting in a total of $2Ns$ segments.

Stage 3.3 Local Trend Computation

In this stage, for each of the $2Ns$ segments, we compute the local trend. To achieve this, we use a least-squares method to fit a polynomial of degree m to the profile of each segment. This polynomial fitting aims to approximate the profile's features as closely as possible. Next, we calculate the variance for each segment using the below two expressions:

For the initial Ns segments:

$$F^2(s,v) = \frac{1}{s}\sum_{j=1}^{s} \{Y[(v-1)s+j] - y_v(j)\} \tag{9}$$

For the latter segments, ranging from $N_s + 1$ to $2N_s$:

$$F^2(s,v) = \frac{1}{s}\sum_{j=1}^{s} [Y((v-N_s)s+j) - y_v(j)]^2 \tag{10}$$

Here, $y_v(j)$ symbolizes the polynomial fit in the segment v. These variance calculations offer foundational data points that will be vital in the subsequent steps of the MF-DFA process.

Stage 3.4 Averaging Over Segments

In this step, we calculate the mean across all segments from the second step to obtain the q^{th} order fluctuation functions. The fluctuation functions, $F_q(s)$, are given as follows:

For $q \neq 0$:

$$F_q(s) = \left[\frac{1}{2Ns}\sum_{v=1}^{2Ns} [F^2(s,v)]^{q/2}\right]^{1/q} \tag{11}$$

For $q = 0$:

$$F_q(s) = \frac{1}{4Ns}\sum_{v=1}^{2Ns} \ln [F^2(s,v)] \tag{12}$$

q aids in differentiating segments based on the size of their fluctuations. A negative q emphasizes smaller fluctuations, while a positive value emphasizes larger ones. Furthermore, when $q = 2$, it aligns with the recognized detrended fluctuation analysis. It is evident that $F_q(s)$ grows as s increases.

Stage 3.5 Scaling Exponent Determination

In the final stage, the scaling exponent of the fluctuation function for a given q is determined. This is done by analyzing the relationship between $F_q(s)$ and s. If $F_q(s)$ follows a power-law behavior, the series will exhibit a linear relationship when plotted on a log-log scale for that particular q.

$$F_q(s) \propto s^{h_q} \tag{13}$$

where h_q is the generalized Hurst exponent.

For stationary series, h_2 corresponds to the standard Hurst exponent. The values of this exponent between 0.5 and 1 suggest positive autocorrelation, denoting a persistent behavior. This is marked effects of long-term memory that manifest irrespective of the time scale. As the values approach 1, they hint at rapid shifts and sudden variations. Conversely, a Hurst exponent between 0 and 0.5 showcases negative autocorrelation, signifying a trend reversal or anti-persistent behavior. Anti-persistence is characterized by covering shorter spans and reverting more regularly than a random event. When $h_q = 0.5$ (with $q=0$), it corresponds to a Brownian series or a traditional random walk. In a broader context, a series demonstrates multifractality when h_q is contingent on q and descends as q ascends. In contrast, the series is mono-fractal when h_q remains unaffected by q.

The above process illustrates the MF-DFA analysis that we conduct on Bitcoin and Altcoins, focusing on their daily and weekly values for the largest and most representative interval with the most recent data to understand their multifractal nature.

3.2.1.2 The Proposed Cryptocurrency Price Forecasting Framework

To Forecast future exchange prices, we propose a forecasting framework for Bitcoin and Altcoins, illustrated in Figure 2. The details of each step in the proposed framework are summarized in Algorithm 2.

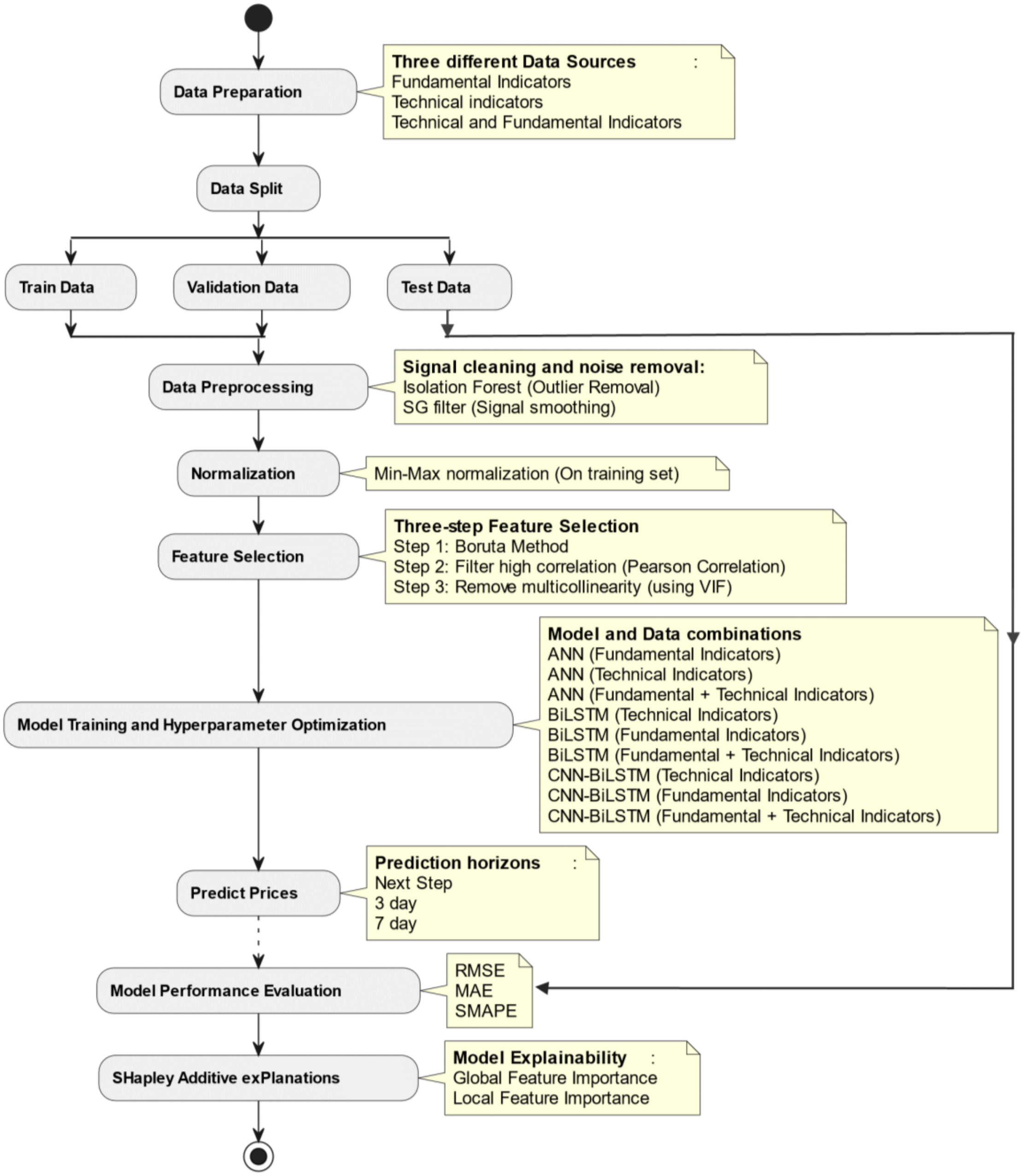


Figure 2. Systematic Workflow of the Proposed Cryptocurrency Price Forecasting Framework

Algorithm 2. Cryptocurrency Price Forecasting Framework

```
Output: Predicted cryptocurrency prices
Input: Daily data from three sources: Fundamental indicators, Technical indicators, and both combined.
for each coin in [Bitcoin, Ethereum, Litecoin, Ripple, Dash, ...] do
    Data Preparation:
    Extract $F = [f_1, f_2, \ldots, f_N]$ (Fundamental)
    Extract $T = [t_1, t_2, \ldots, t_N]$ (Technical)
    Extract $C = [c_1, c_2, \ldots, c_N]$ (Combined)
    Data Split:
    Partition data into training (70%), validation (15%), and testing (15%)
    Data Preprocessing:
    Apply Isolation Forest to train data and remove outliers
    Suppress signal noise using Savitzky-Golay Filter:
    For each point $i$ in the data sequence:
    $y'_i = \sum_{j=-\frac{m-1}{2}}^{\frac{m-1}{2}} c_j y_{i+j}$
Where: $y'_i$ is the smoothed value at point $i$, $m$ is the window size of the filter, $c_j$ are the filter coefficients, $y_{i+j}$ are
    the data points in the window
    Normalization:
    Apply Min-Max normalization:
    $X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$
Where: $X$ is the original data value, $X_{min}$ and $X_{max}$ are the minimum and maximum values in the dataset,
    respectively
    Feature Selection:
    Use Boruta Method: Retain feature if $Z - Score_{original} > Z - Score_{Max\,shadow}$
Where: $Z - Score_{original}$ is the importance score of the original feature, $Z - Score_{Max\,shadow}$ is the highest
    importance score among shadow features, Shadow features are randomized copies of the original features
    Pearson Correlation: $r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$
Where: $r$ is the correlation coefficient, $x_i$ and $y_i$ are data values, and $\bar{x}$ and $\bar{y}$ are their means
    VIF: $VIF_i = \frac{1}{1 - R_i^2}$
Where: $VIF_i$ is the Variance Inflation Factor for the $i^{th}$ variable, $R_i^2$ is the coefficient of determination
    Model Training:
    ANN: $f_{ANN}(X; \theta_{ANN})$
    BiLSTM: $f_{BiLSTM}(X; \theta_{BiLSTM})$
    CNN-BiLSTM: $f_{CNN-BiLSTM}(X; \theta_{CNN-BiLSTM})$
Where: $X$ is the set of input features, $\theta_{ANN}, \theta_{BiLSTM}, \theta_{CNN-BiLSTM}$ are the hyperparameters to be optimized
    using Bayesian Optimization for the respective models
    Predict Prices:
    For each model $M$ and input set $X$:
    Forecast for next step: $\hat{Y}_{1-step} = M(X; \theta)$
    Forecast for 3 steps ahead: $\hat{Y}_{3-step} = M(X; \theta)$
    Forecast for 7 steps ahead: $\hat{Y}_{7-step} = M(X; \theta)$
Where: $M$ is the trained model (ANN, BiLSTM, or CNN-BiLSTM),$X$ is the set of input features, $\theta$ is the set of
    optimized parameters for $M$, $\hat{Y}_{k-step}$ is the forecasted price $k$ steps ahead
    Model Evaluation:
    RMSE: $\sqrt{\frac{1}{N} \sum (Y_i - \hat{Y}_i)^2}$
Where: $Y_i$ is the actual value, $\hat{Y}_i$ is the predicted value, and $N$ is the total number of values
    MAE: $\frac{1}{N} \sum |Y_i - \hat{Y}_i|$
    SMAPE: $\frac{100\%}{N} \sum \frac{|Y_i - \hat{Y}_i|}{|Y_i| + |\hat{Y}_i|}$
    SHAP Analysis:
    Evaluate global and local feature importances
end for
```

We start the process by extracting daily data from three distinct sources: fundamental indicators, technical indicators, and a combination of both. We first clean and normalize the data. Given the high dimensional multivariate nature of the dataset, we perform min-max scaling to ensure consistency. After normalizing the data, we split the dataset into training, validation, and testing sets at a 70:15:15 ratio. We train the model on the training set, use the validation set for hyperparameter optimization, and evaluate its performance on the test set, treating it as unseen out-of-sample data.

After partitioning the data, we initiate the preprocessing steps. The inherent high-frequency fluctuations and noise in cryptocurrency data, largely driven by market volatility, make it susceptible to potential outliers. We utilize the Isolation Forest algorithm to identify and manage these outliers systematically (F. T. Liu et al., 2012). Isolation Forest isolates observations by selecting a random feature and then partitioning the data between its minimum and maximum values. The Isolation Forest identifies anomalies by counting the splits to isolate each point. Points that require fewer splits are more likely considered anomalies. Careful parameter calibration ensures optimal results when employing the Isolation Forest for outlier detection in cryptocurrency data. Genuine outliers, indicative of legitimate market shifts, should not be inadvertently removed. We ensure this by adjusting. We adjust the model to handle varying data sample sizes and control the proportion of outliers the model expects. The model's flexibility allows us to selectively choose a subset of features for training, potentially improving anomaly detection. Additionally, we determine the sampling strategy, influencing the randomness and diversity of data subsets used in training. We also prioritize reproducibility by pseudo-randomness of the feature selection, ensuring that our detection remains consistent across multiple evaluations. This fine-tuning ensures that genuine market shifts are recognized and retained while still robustly identifying actual outliers.

After handling outliers, we address the high-frequency noise fluctuations in the cryptocurrency data. We utilize the Savitzky–Golay Filters with Fitting Weights (SGW) for this. By fitting successive subsets of adjacent data points with a dynamically chosen low-degree polynomial, the SGW effectively smoothens the time series, preserving critical price trends and preventing models from being influenced by misleading, minor fluctuations.

For feature selection, we employ a three-step process. First, we utilize the Boruta method to rank features based on their importance compared to random shadow features, ensuring only the most significant predictors are retained. Subsequently, we apply Pearson Correlation to identify linear relationships between individual pairs of features. By doing so, we identify and eliminate features that exhibit high correlations with others, potentially indicating redundancy. However, addressing pairwise correlations alone does not fully address multicollinearity, wherein three or more variables exhibit high correlation. Pearson's method does not capture such multicollinearity. We implement the Variance Inflation Factor (VIF) as the final step to mitigate this. VIF quantifies how much the variance of an estimated regression coefficient increases when predictors are correlated. If certain features collectively predict another feature beyond a set threshold, often defined by a VIF of 10, we deem it indicative of problematic collinearity and remove the offending variables. This process ensures our model's stability, interpretability, and resistance to overfitting that might arise due to multicollinearity.

During the model training phase, we employ three distinct machine learning architectures, each with strengths and weaknesses: the Artificial Neural Network (ANN), known for capturing non-linear relationships. The Bidirectional Long Short-Term Memory (BiLSTM) is optimized for handling time series data, and the hybrid Convolutional Neural Network-Bidirectional Long Short-Term Memory (CNN-BiLSTM) combines spatial feature extraction with sequence modeling. Through these diverse architectures, we aim to identify which model best fits our data. We train these models using three data combinations: solely on fundamental indicators, exclusively on technical indicators, or a combination of both. We then use Bayesian Optimization on the validation set for hyperparameter tuning.

The models are then used to predict future prices at various horizons. We then evaluate the performance of these models using metrics such as RMSE, MAE, and SMAPE for each of the four intervals. We repeat this process for each coin namely Bitcoin, Ethereum, Litecoin, Ripple, and Dash.

The subsequent sub-sections provide the experimental settings using in the framework with detailed theoretical background for the methodologies used in the proposed financial forecasting framework.

3.2.1.3 Experimental Setup and Theoretical Background of the Proposed Framework

3.2.1.3.1 Data Extraction and Preparation

The proposed forecasting framework necessitates a robust dataset to aptly represent the dynamic nature of the cryptocurrency market. To effectively represent the cryptocurrency market, our data is categorized into three primary domains. First, the fundamental indicators delve into the foundational value of cryptocurrencies. They encompass macroeconomic variables such as global currency to USD ratios, oil and gold prices, and the Federal treasure rate; speculative indicators derived from sources like Google Trends and Twitter data; and blockchain-specific variables including mining information, network activity, wallet activity, transaction details, wealth distribution, and market signals. These combined factors give a robust understanding of the inherent value and external perceptions surrounding cryptocurrencies. Technical indicators, which are calculations based on historical price, volume, momentum, and oscillation patterns, like Moving Averages and Relative Strength Index (RSI), play a pivotal role. They help to capture the cyclical patterns and price movement trends inherent in financial markets. Recognizing the strengths of both domains, we also create a combined set of indicators. By merging technical and fundamental data, this dataset offers a comprehensive view, considering both past price trends and core cryptocurrency metrics. Tables 1 and 2 provide a detailed list of all variables used in each of the three categories of input variables used in our multivariate analysis.

The partitioning of data is important step in achieving accurate model predictions. We designate 70% of the data to the training set, enabling the model to discern and adapt to underlying patterns in the data. The validation set, comprising 15% of the data, aids in model optimization through hyperparameter tuning, ensuring the model's robustness without leaning on the test set. The remaining 15% is set aside as the testing set, serving as an evaluative benchmark, assessing the model's forecasting capability on unseen data, and providing insights into its prospective real-world performance. In our study, the decision to use a 70:15:15 split for dividing the data into training, validation, and testing sets was guided by several considerations. Firstly, when dealing with a limited quantity of data, a 70:15:15 split allows for retaining a larger portion of the data in the training set while still providing reasonably sized validation and testing sets. This approach helps to balance the model’s exposure to training data with the need for effective model validation and testing. Secondly, a 70:15:15 split can reduce the chances of overfitting. By not overly skewing the dataset towards training (as in an 80:10:10 split), we mitigate the risk of the model becoming too closely fitted to the training data, which could compromise its performance on new, unseen data.

To ensure that our models are consistent and effectively capture the varying dynamics of each specific cryptocurrency, we segment our data into distinct intervals. Each interval corresponds to a distinct phase in the cryptocurrency market's development. The first interval reflects the initial stages and development of the cryptocurrency market. The second interval is significant as it captures the market's trajectory post the Mt. Gox crash, highlighting phases of market stabilization and responses to regulatory changes. The third interval is particularly notable due to the significant price surges experienced, especially in December 2017 and November 2021. The fourth interval spans the post-2017 era, during which cryptocurrencies received increased recognition from mainstream financial institutions, amidst varying price corrections. Tables 3 through 7 present the specific durations associated with these intervals for each cryptocurrency, including Bitcoin, Ethereum, Litecoin, Ripple, and Dash.

3.2.1.3.2 Data Normalization and Signal Processing

Given the multivariate nature of our dataset, with each variable potentially having different characteristics and scales, normalization becomes indispensable. Min-Max scaling is chosen for its simplicity and efficiency in bringing all variables to a common scale without distorting the differences in the range of values. By transforming each feature to a predefined range, usually [0,1], Min-Max scaling ensures that each feature influences the model based on its intrinsic importance rather than its original scale.

3.2.1.3.3 Outlier Removal using Isolation Forest Algorithm

Cryptocurrency prices, given their volatile nature, can exhibit sudden spikes or drops. These fluctuations can be genuine market reactions or anomalies. The Isolation Forest algorithm (F. T. Liu et al., 2008) is used to differentiate between these suspected outliers and regular data points. A group of Isolation Trees composes an Isolation Forest. Each tree is constructed from a random subset of the data, and anomalies are identified using the average path lengths across all trees. The process starts with the recursive partitioning of the dataset to create an 'Isolation Tree.' In each node of this tree, a feature is randomly selected, and a split value is determined between the maximum and minimum values of the chosen feature. The dataset is then divided based on this split value, and this bifurcation continues until data points are fully isolated or a predetermined tree depth is reached. The path length is an essential indicator which measures the depth at which an instance becomes isolated. Anomalies, by their distinct characteristics compared to most data points, tend to be isolated at earlier stages in the tree, resulting in shorter path lengths. The Isolation Forest algorithm's strength lies in its ensemble approach. The algorithm ensures robust and consistent anomaly detection by constructing multiple Isolation Trees, each from a random subset of the dataset, and then averaging the path lengths across all trees. The Isolation Forest algorithm offers a systematic approach to identifying outliers while preserving genuine market shifts. Additionally, Isolation Forest can handle high-dimensional datasets efficiently due to its ability to randomly select features for splitting at each tree node. To implement the Isolation Forest algorithm effectively, the following hyperparameters were meticulously calibrated, considering our dataset's characteristics:

i. *n_estimators*: Number of trees in the Isolation Forest (Range: 100 to 500, step: 50)
ii. *max_samples*: Number of samples to draw from X to train each base estimator (Range: 100 to 300, step: 20)
iii. *Contamination*: Proportion of outliers in the data set (Range: 0.05, 0.1, 0.15, 0.2)
iv. *max_features*: Number of features to draw from raw dataset to train each base estimator (Range: 5, 10, 15, 20)
v. *bootstrap*: Whether samples are drawn with replacement (Values: True, False)
vi. *n_jobs*: Number of jobs to run in parallel for both fit and predict (Values: 1, 5, 10)

These parameters were set with the aim of optimizing the algorithm's performance in accurately identifying outliers while handling the high-dimensional nature of the dataset efficiently. The Isolation Forest algorithm offers a systematic approach to outlier detection by averaging the path lengths across all trees in the ensemble.

The Isolation Forest algorithm assigns each data point an "anomaly score" that indicates how likely it is to be an outlier. The higher the score, the more unusual the data point is relative to the rest of the dataset. This score is crucial for determining which points are considered anomalies. Mathematically, the score for a sample x is given by:

$$s(x, n) = 2 - \frac{E(h(x))}{c(n)} \tag{14}$$

where $E\big(h(x)\big)$, represents the path length, indicating the number of edges an observation must pass in the tree to be isolated. This value captures the average depth a point travels in the tree across all trees in the forest.

The equation for this path length is:

$$E(h(x)) = 2H(n-1) - \frac{2(n-1)}{n} \tag{15}$$

where, $H(i)$ is the harmonic number, and in the context of Isolation Forest, it denotes the expected value or average depth of the tree; $c(n)$ serves as a normalization factor, derived from the average path length of an unsuccessful search in a binary search tree.

By normalizing, the anomaly score s is ensured to lie between 0 (indicating typical data points) and 1 (indicating potential anomalies). When a data point in an Isolation Tree has a short path length, it indicates that fewer splits are required to isolate it, indicating its anomalous nature. Therefore, anomalies will have shorter path lengths on average than regular data points in the ensemble of trees that comprise the Isolation Forest. The Isolation Forest algorithm utilizes this differentiating characteristic, the average path length, to classify data points as either anomalies or typical observations.

3.2.1.3.4 Signal Smoothing using Savitzky-Golay Filters with Fitting Weights

After removing the outliers from the multivariate time series data of Bitcoin and various Altcoins using the Isolation Forest algorithm, we turn our attention to suppressing the noise and enhancing the clarity of the data. We employ the Savitzky-Golay Filters with Fitting Weights (SGW) method.

The Savitzky-Golay filter, a type of smoothing filter, enhances the accuracy of input series by reducing noise while preserving essential signal characteristics such as amplitude, frequency, and phase components, and broad trends (Schafer, 2011). The essence of the filter is its capability to fit a low-degree polynomial (often linear or quadratic) to a series of sequential data points. This polynomial's coefficients are determined by a convolution process, which involves the original data points and a set of convolution coefficients.

Savitzky-Golay filters operate on the principle of local polynomial regression (LPR) (Ledolter, 2013). For a given noisy signal represented as $y_n = x_n + w_n$, where y_n is the observed noisy value at index n, the objective is to approximate a local subset of the data using a polynomial of order N and then update each data point with the value of the polynomial at that point. Mathematically, considering a symmetric window over indices $n = -M, \ldots, M$, the polynomial approximation at any index n is given by (Schafer, 2011):

$$p(n) = \sum_{k=0}^{N} a_k n^k \tag{16}$$

The goal is to determine the polynomial coefficients a_k that minimize the mean squared error between the actual and the fitted values. This error (cost) is formulated as:

$$\mathcal{E} = \sum_{n=-M}^{M} \left(\sum_{k=0}^{N} a_k n^k - x[n]\right)^2 \tag{17}$$

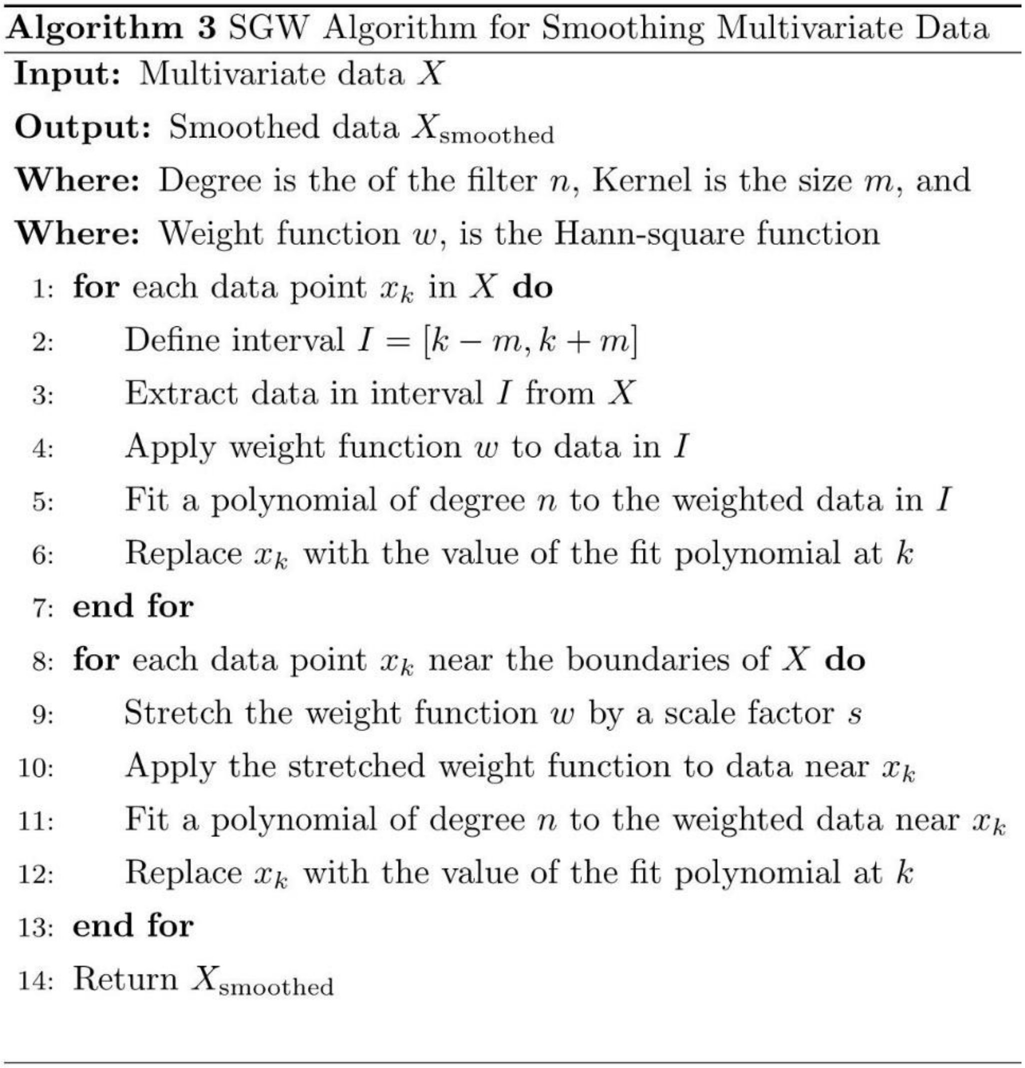
The Savitzky-Golay (SG) Filters with Fitting Weights (SGW) is an enhancement over the traditional SG method. It introduces weights during polynomial fitting, mitigating the influence of periphery data points within the window. Mathematically, the polynomial fitting in the SGW method can be represented as (Schmid et al., 2022):

$$p_w(n) = \sum_{k=0}^{N} a_k n^k w(n) \tag{18}$$

The introduction of this weight function means that the error formulation also needs to account for these weights. The error in the SGW method becomes:

$$E_w = \sum_{n=-M}^{M} \left(\sum_{k=0}^{N} a_k n^k w(n) - x[n]\right)^2 \tag{19}$$

By minimizing the weighted error, the SGW filter optimally fits polynomial coefficients, considering the significance of each data point within the window. The procedure is described step-by-step in Algorithm 3 using the method suggested by (Schmid et al., 2022):



Algorithm 3 SGW Algorithm for Smoothing Multivariate Data

```
Input: Multivariate data $X$
Output: Smoothed data $X_{\text{smoothed}}$
Where: Degree is the of the filter $n$, Kernel is the size $m$, and
Where: Weight function $w$, is the Hann-square function
for each data point $x_k$ in $X$ do
    Define interval $I = [k - m, k + m]$
    Extract data in interval $I$ from $X$
    Apply weight function $w$ to data in $I$
    Fit a polynomial of degree $n$ to the weighted data in $I$
    Replace $x_k$ with the value of the fit polynomial at $k$
end for
for each data point $x_k$ near the boundaries of $X$ do
    Stretch the weight function $w$ by a scale factor $s$
    Apply the stretched weight function to data near $x_k$
    Fit a polynomial of degree $n$ to the weighted data near $x_k$
    Replace $x_k$ with the value of the fit polynomial at $k$
end for
Return $X_{\text{smoothed}}$
```

We set the filter's degree, n, to denote the polynomial's order and define the kernel size, m, to establish the data window's breadth. The Hann-square function, commonly known as cos^4, is our weight function due to its prowess in sidelobe suppression and seamless transitions. We designate a local interval for every data point, weight the data within it, and fit a polynomial of degree n. Data points near the edges receive heightened attention, stretching the weight function to ensure uniform smoothing. Consequently, the SGW filter yields a signal with enhanced noise reduction and precise depiction of features, notably around dataset boundaries. By applying SGW method we aim to attenuate noise within the cryptocurrency time series while preserving intrinsic signals and minimizing the presence of artifacts.

3.2.1.3.5 Feature Selection

After preprocessing and noise suppression in our multivariate time series data of cryptocurrencies, the next important step is selecting the most significant features that contribute effectively to the prediction model. Feature selection not only reduces the dimensionality, making the training process more efficient, but also helps in avoiding overfitting and enhances the generalization capabilities of the model.

Step1: Feature Selection Using Z-Score with the Boruta Method

The Boruta algorithm is a wrapper around a random forest classifier that identifies and retains those features that are decisively more important than random permutations of the dataset (Kursa & Rudnicki, 2010). It differentiates between informative and non-informative features by comparing their importance against randomized copies of these features, termed shadow features. Shadow features are generated by randomly permuting the values of the original features, effectively removing any inherent correlation with the target variable. A random forest algorithm, including these shadow features, is trained on the dataset. Mathematically, for each original feature, the importance score, denoted by the Z-Score, is compared to the highest importance score among its shadow features (randomized copies).

A feature is retained if:

$$Z_{\text{Score(original)}} > Z_{\text{Score(Max shadow)}} \tag{20}$$

This comparison is iterative. In each iteration, non-essential features are pruned, and the process is repeated until a definitive decision is made for all features. The strength of the Boruta algorithm lies in its systematic comparison against shadow features, ensuring that only those features which demonstrate importance significantly above random chance are retained.

Step 2: Eliminating Redundant Features Using Pearson Correlation

To evaluate the linear relationship between each feature and the cryptocurrency prices, we employ Pearson's correlation coefficient. A practical method for selecting important features involves filtering out those with a high correlation coefficient (D. Kumar et al., 2016). This coefficient measures the linear correlation between two datasets and can range between -1 and 1. Features with correlation values close to 0 exhibit weak linear relationships with the target. Through rigorous experimentation and aligning with findings from established studies in our domain (Mudassir et al., 2020; Tripathi & Sharma, 2022), we determine a threshold value of 0.80 for the Pearson correlation coefficient. We exclude features with a correlation value below this threshold, ensuring we retain only the most influential variables for our analysis, we filter out features with a correlation value less than 0.80, as these tend to be less impactful for our predictions. We compute the Pearson Correlation for any two variables as:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{21}$$

Step 3: Removing Multicollinearity by Filtering High VIF

Even after employing the Boruta method and eliminating highly correlated features, residual multicollinearity may persist within the dataset. To address this, we leverage the Variance

Inflation Factor (VIF). VIF serves as a robust metric that quantifies the degree of multicollinearity present. When independent variables exhibit high inter-correlations, it can compromise the reliability of the model's estimates, often resulting in unstable and inaccurate predictions. For each feature, the VIF is calculated using the coefficient of determination, R2, as per the following formula:

$$VIF_i = \frac{1}{1-R_i^2} \quad (22)$$

A feature with a VIF exceeding our set threshold of 10 indicates that it shares significant linear relations with other features (Jang & Lee, 2018b). Such high VIF values can distort the predictive model's performance. In our analysis, the Variance Inflation Factor (VIF) range is set between 1 and 10. A VIF above 10 signifies significant linear relationships between features, impacting the model's accuracy. By monitoring and setting a strict VIF range threshold, we aim to ensure that our dataset is finely tuned for optimal modeling, minimizing the impact of multicollinearity on our predictions.

3.2.1.3.6 Model Training and Hyperparameter Optimization

Deep learning models, with their innate capacity to recognize and exploit intricate patterns within datasets, play a pivotal role in our cryptocurrency forecasting framework. As we endeavor to achieve precise and reliable price predictions, we employ a suite of deep learning architectures, each bringing its unique strengths to the table. The models are trained on three distinct input sets: fundamental indicators, technical indicators, and a composite of both.

i) Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are computational models inspired by the human brain. ANNs are composed of layers of interconnected neurons whose weights and biases are adjusted during training. This structure enables them to act as universal approximators, representing complex, nonlinear relationships in various tasks. The primary components of ANNs are neurons organized into layers: an input layer, multiple hidden layers, and an output layer. The input layer is the initial layer where the data enters the network. It transforms the raw input into a format suitable for subsequent layers. Hidden Layers are layers between the input and output layers. Each neuron in a hidden layer receives input from all neurons of the previous layer, processes it with a weighted sum followed by an activation function, and passes the output to every neuron in the next layer. Output Layer is the final layer, which produces the result for given inputs.

The transformation of data between any two consecutive layers can be mathematically represented as (Ghasemiyeh et al., 2017):

$$a^{(l)} = \sigma\left(\mathrm{W}^{(l)} a^{(l-1)} + \mathrm{b}^{(l)}\right) \text{ for } l = 1,2,\dots,L \quad (23)$$

where $a^{(l)}$ is the activation function of the l^{th} layer, $\mathrm{W}^{(l)}$ is the weight matrix for layer l which connects it to layer $l-1$, $\mathrm{b}^{(l)}$ is the bias vector for the l^{th} layer, σ is the activation function applied element-wise to its argument.

$$y = \sigma\left(\sum_{l=1}^{p} w_{o,l}^{(L+1)} a_l^{(L)} + b_o^{(L+1)}\right) \tag{24}$$

where y is the predicted value from the output layer;
σ is the activation function used in the output layer;
L denotes the number of hidden layers in the network;
$w_{o,l}^{(L+1)}$ represents the weight connecting the l^{th} neuron from the final hidden layer to the output neuron;
$a_l^{(L)}$ is the activation of the l^{th} neuron from the final hidden layer;
$b_o^{(L+1)}$ is the bias term associated with the output neuron;
p is the total number of neurons in the final hidden layer.

Figure 3 illustrates the schematic structure of an ANN, which includes an input layer, three hidden layers, and an output layer.

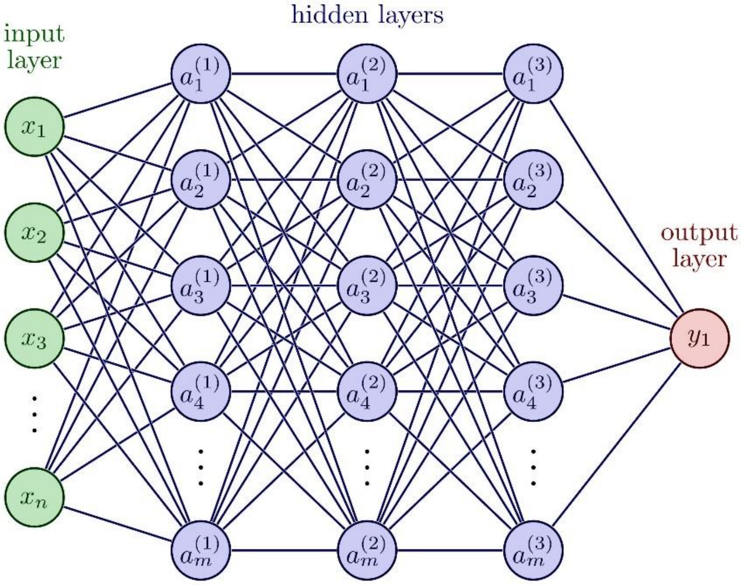


Figure 3. Schematic structure of an Artificial Neural Network

In our study, we design three distinct ANN models, each tailored to a specific type of indicator: technical, fundamental, and composite. This segmentation ensures that each model is finely tuned to the unique characteristics, specific relationships, and patterns inherent within each type of indicator. The customized approach enhances the capacity of the suggested framework to identify and analyze the patterns inherent in the underlying data.

Furthermore, we employ Hyperparameter Optimization (HPO) using Bayesian Optimization to fine-tune the network. We systematically optimize a range of hyperparameters, including the number of neurons in each layer, the number of hidden layers, and the choice of activation functions. These optimizations are carried out using HPO to ensure optimal network performance.

ii) Bidirectional Long Short-Term Memory (BiLSTM)

Time series data, such as cryptocurrency prices, inherently possess temporal dependencies that span over varying durations. LSTMs, a specialized form of recurrent neural networks that are equipped to remember and utilize these long-term dependencies. LSTM units consist of a cell, an input gate, an output gate, and a forget gate. The cell remembers values over arbitrary time intervals, and the three gates regulate the flow of information into and out of the cell. The LSTM by itself is unidirectional, meaning that at each time step of the sequence, it is processing data from the past to the future (from left to right). The main gates in an LSTM are the forget gate f_t, input gate i_t, and output gate o_t. Additionally, the cell state, C_t, is updated using the values from these gates. The equations governing these components can be summarized as follows (Kavianpour et al., 2021):

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{25}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{26}$$

$$\tilde{C}_t = \tanh\,(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{27}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{28}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{29}$$

$$h_t = o_t \times \tanh\,(C_t) \tag{30}$$

where f_t is the forget gate's activation at time t, i_t and o_t are the input and output gate's activations, respectively. $\tilde{C}_t$ is the cell's candidate value, C_t is the cell's final value, ht is the output, σ is the sigmoid activation function, W and b are weights and biases, respectively, with appropriate subscripts denoting their use, and x_t is the input at time t.

The bidirectional iteration of LSTM further refines this capability. In a Bidirectional LSTM (BiLSTM), two separate memory cells are used for forward and backward passes through the sequence. The outputs of the forward and backward LSTMs for each time step can either be summed or concatenated, which gives the final output of the BiLSTM for that time step.

$$\overrightarrow{h_t} = \overrightarrow{LSTM}(x_t, \overrightarrow{h_{t-1}}) \tag{31}$$

$$\overleftarrow{h_t} = \overleftarrow{LSTM}\left(x_t, \overleftarrow{h_{t+1}}\right) \tag{32}$$

where $\overrightarrow{LSTM}$ and $\overleftarrow{LSTM}$ are the forward and backward LSTM functions, respectively.

The final hidden state h_t for the BiLSTM at time t is typically obtained by concatenating the forward and backward hidden states:

$$h_t = \begin{bmatrix} \overrightarrow{h_t} \\ \overleftarrow{h_t} \end{bmatrix} \tag{33}$$

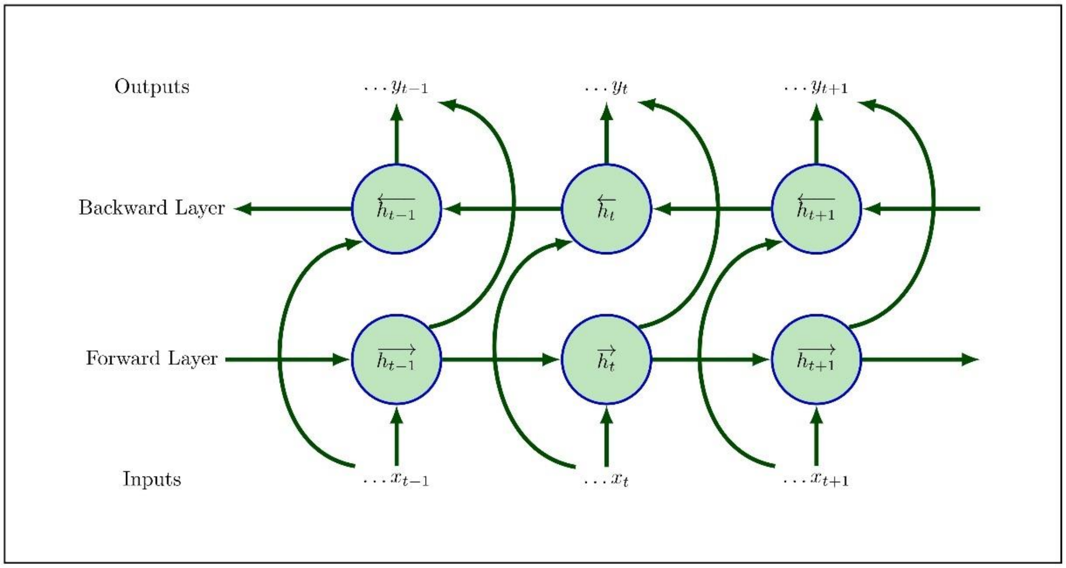


Figure 4. Schematic diagram of a BiLSTM Network

Figure 4 illustrates the forward and backward passes of a BiLSTM network. In practice, the forward and backward passes of the BiLSTM can be computed independently, and their outputs are combined only at the end. This combined representation ensures that the information from both past (left) and future (right) of a sequence element is considered for any given time step.

We apply self-attention to enhance the BiLSTM network's ability to capture and emphasize crucial long-range dependencies within the sequence of technical, fundamental, and composite indicators. The self-attention mechanism provides a weighted representation of the sequence by allowing the model to focus on different parts of the input based on their relevance. The mathematical formulation of the self-attention mechanism is given as follows (Kavianpour et al., 2021):

Firstly, for each time step in our BiLSTM output, we compute three vectors: the query Q, the key K, and the value V. These vectors are derived by multiplying the BiLSTM output with weight matrices:

$$Q = hW_Q, K = hW_K, V = hW_V \tag{34}$$

where h represents the BiLSTM's output, W_Q, W_K, and W_V are the weight matrices for queries, keys, and values.

In the next step, we apply the scaled dot product attention. For each query vector q_t, we compute attention scores against all key vectors. This score gives a measure of relevance or "attention" a specific key-value pair should receive:

$$\text{scores}\ (t, j) = \frac{q_t \cdot k_j}{\sqrt{d_k}} \tag{35}$$

with d_k being the dimensionality of the queries and keys, t is the current time step, j represents the other positions in the sequence that are compared to the current time step t to compute attention scores.

The next step in the self-attention process is Softmax Normalization. We apply the softmax function to the scores from equation (35), converting them into a probability distribution that ensures values range between 0 and 1.

$$\alpha_{t,j} = \text{softmax}\ (\text{scores}\ (t, j)) \tag{36}$$

The subsequent step involves calculating the attention weights, which allows us to generate a weighted sum of the value vectors:

$$a_t = \sum_{j=1}^{T} \alpha_{t,j} v_j \tag{37}$$

where a_t is the attention-enhanced output for the sequence at position t

Integrating the self-attention mechanism enhances the BiLSTM model's capability to simultaneously capture local context from the BiLSTM layers and the global context of the entire sequence, optimizing its efficacy in comprehending intricate temporal dependencies.

iii) CNN-BiLSTM

Integrating Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) models presents a holistic methodology for the analysis and understanding of data. CNNs excel at analyzing data to extract localized patterns or features, whereas BiLSTMs provide temporal context to these patterns. Figure 5 shows the architecture of a CNN-BiLSTM with self-attention mechanism network.

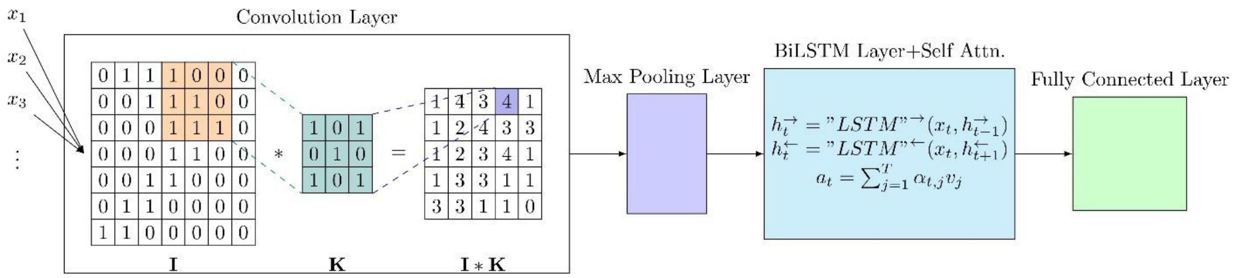


Figure 5. Schematic Architecture of a CNN-BiLSTM Network with Self Attention

The CNN-BiLSTM architecture utilizes early convolutional layers to identify important data features efficiently. The BiLSTM layers are then given these identified features for a complete temporal analysis. The CNN-BiLSTM model combines the strengths of both architectures: CNNs excel at detecting distinct patterns within data, while BiLSTMs analyze these patterns concerning their sequence. In the CNN-BiLSTM model, initial convolutional layers detect important features from the data. These features are then passed to the BiLSTM layers for temporal analysis. The BiLSTM layers use the Self-Attention mechanism which allows the model to weigh the importance of these features differently based on their context in the sequence. This combination ensures the model can recognize and understand important patterns spatially, over time, and in relation to other features. This combination ensures the model can recognize and understand important patterns spatially and over time. The Self-Attention mechanism can be described mathematically as:

$$\text{Attention }(Q, K, V) = \text{Softmax }\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{38}$$

where Q, K, and V are the query, key, and value matrices, respectively, and d_k is the dimension of the keys.

The convolution operation in the layer is described as:

$$X_i^l = \sigma\left(W_i^l * X^{l-1} + b_i^l\right) \tag{39}$$

where X_i^l is the feature map of the i^{th} filter in the l^{th} layer, W_i^l is the i^{th} filter in the l^{th} layer, X^{l-1} is the input from the previous layer, b_i^l is the bias of the i^{th} filter in the l^{th} layer, and σ is the activation function.

The pooling operation is given by:

$$Y_i^l = \text{pool }\left(X_i^l\right) \tag{40}$$

where Y_i^l is the output after pooling, and pool represents the pooling function (like max-pooling).

The operation in the fully connected layer can be described as:

$$Z^l = \sigma(W^l \cdot Y^{l-1} + b^l) \tag{41}$$

where Z^l is the output of the l^{th} layer, W^l and b^l are the weight and bias of the l^{th} layer respectively, Y^{l-1} is the input from the previous layer, and σ is the activation function.

To ensure these models deliver their optimal performance, we employ Bayesian optimization for hyperparameter tuning. Bayesian optimization systematically explores the hyperparameter space, evaluating regions that show promise and progressively converging towards optimal model configurations. By leveraging Bayesian optimization, our objective is to meticulously fine-tune the parameters of our models, setting them up to achieve their utmost potential and, in turn, ensuring precise and robust cryptocurrency price predictions.

In our modeling process, various hyperparameters such as the learning rate, dropout rate, number of hidden layers, and kernel size, among others, play an important role. A comprehensive list of all these parameters and their configurations can be found in Table 10.

iv) Hyperparameter Optimization using Bayesian Optimization

Hyperparameter optimization (HPO) is fundamental in constructing efficient and effective neural network models (Zulfiqar et al., 2022). In deep learning, where models vary from feedforward architectures in Artificial Neural Networks (ANN) to backpropagation-based sequences in Long Short-Term Memory (LSTM) and convolution operations in CNNs, the right choice of hyperparameters can substantially influence a model's performance. External to the training process, these hyperparameters guide the model's learning behavior, making their optimal selection critical.

Bayesian Optimization (BO) provides a systematic approach to Hyperparameter Optimization (HPO), distinguishing itself from conventional techniques like grid and random search (Ranjit et al., 2019). Instead of exhaustively testing combinations as in grid search or relying on random sampling like in random search, BO strategically chooses hyperparameters based on prior evaluations. It uses a probabilistic model, typically a Gaussian Process, to predict the performance of a new set of hyperparameters without directly evaluating them, ensuring a balance between exploring new regions and exploiting known good regions. In this work, we use Facebook's (now Meta) research team's Adaptive Experimentation Platform Library in Python for hyperparameter optimization of various deep learning models used in our framework (Bakshy et al., 2018). This platform is designed to effectively handle evaluations affected by noise; a common difficulty faced while training deep learning models. Additionally, it provides multi-objective optimization, enabling the concurrent optimization of numerous objectives that may conflict with one another. We use the same hyperparameter search space listed in Table 10 for all the experiments in our proposed financial forecasting framework.

Table 10. Hyperparameter search space for Bayesian Optimization in deep learning model tuning

Hyperparameter Name	**Type**	**Bounds**	**Hyperparameter Value Range**	**Description**
Learning rate	range	{0.001, 0.01}	-	The learning rate for network weights
Dropout rate	range	{0, 0.25}	-	Regularization dropout to ignore subset of neurons
Count of hidden layers	range	{1, 5}	-	Number of hidden layers
Neuron count per layer	range	{350, 550}	-	Neurons in each layer
Batch size	choice	-	{4, 8, 16, 32, 64, 128}	Batch Size for a single pass
Activation function	choice	-	{linear, relu, sigmoid, tanh, softmax, elu, selu}	Activation function for the network

Optimizer	choice	-	{rms, adam, adamx, sgd, Nadam}	Optimizer for reducing the loss function
CNN filters	choice	-	{32, 64, 128, 256}	Number of convolutional filters
Kernel size	choice	-	{2, 3, 4, 5}	Kernel size for convolutional layers
Pooling size and type	choice	-	{2 to 5} and {‘max’, ‘average’}	Pooling size and type
Gradient Clipping	range	e.g., {0.1, 10}	-	Limits the values of gradients
Weight Decay (L2 reg.)	range	e.g., {0.0001, 0.01}	-	Penalty on the loss for large weights

Note: The range of hyperparameter values presented in this table are based on the author's own experience and findings from past studies.

3.2.1.3.7 Model Performance Evaluation

The forecasting horizons refer to the number of time steps ahead for which predictions are made. In this study, the forecasting horizons pertain to the number of time steps ahead for predictions. We evaluate the model's predictive capabilities across various horizons, including the immediate next step prediction, a 3-step prediction horizon forecasting values three-time steps ahead, and a 7-step prediction horizon anticipating values a week into the future. The evaluation process is repeated for each cryptocurrency to ensure a comprehensive analysis. Given the unique market dynamics, trading volumes, and investor sentiments associated with each coin, individual analyses are paramount. Specifically, we delve into the intricacies of five major cryptocurrencies: Bitcoin, Ethereum, Litecoin, Ripple, and Dash. This approach underscores the importance of understanding the distinct behaviors and patterns inherent to each cryptocurrency. For each model and cryptocurrency, we conduct 15 experimental runs to account for the inherent variability in deep learning training processes. The results from these runs are then averaged, and the mean performance metrics are reported. This approach ensures a more robust and reliable representation of each model's performance.

The choice to conduct 15 experimental runs for each model and cryptocurrency is grounded in the need to account for the inherent variability in deep learning training processes. Deep learning models, especially those used in financial forecasting, are sensitive to initial conditions and data sampling. By averaging the results of multiple runs, we ensure that the reported performance metrics are not skewed by outliers or specificities of a single training instance. This number of runs strikes a balance between computational feasibility and statistical robustness, providing a reliable representation of each model's performance across different scenarios and cryptocurrencies.

We employed three performance evaluation metrics to assess the models used in the proposed framework.

i) Root Mean Square Error (RMSE)

The root-mean-square error (RMSE) measures the differences between the predicted and observed values. By squaring the differences and then taking the square root of the average, RMSE gives higher weight to large errors. This means that RMSE is sensitive to outliers and indicates the magnitude of errors made by the model. Mathematically, RMSE is given as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{42}$$

where y_i is the actual value and $\hat{y}$ is the predicted value.

ii) Mean Absolute Error (MAE)

MAE measures the average magnitude of errors between predicted and observed values, without considering their direction. It calculates the average of the absolute differences between the predicted and actual values, offering a straightforward depiction of prediction accuracy. It is expressed as follows:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{43}$$

where y_i is the actual value and $\hat{y}$ is the predicted value.

iii) Mean Absolute Percentage Error (MAPE)

MAPE is a percentage error metric that measures the size of the error in terms of percentage. It provides a normalized measurement to understand the magnitude of forecast errors in relative terms. MAPE is useful in contexts where it is essential to understand the scale of prediction errors in percentage terms. It is given by:

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{44}$$

where y_i is the actual value and $\hat{y}$ is the predicted value.

When evaluating the model performance, we employ a combination of RMSE, MAE, and MAPE indicators to provide a holistic assessment. RMSE emphasizes the impact of large errors, which is crucial in financial forecasting where significant deviations can be detrimental. MAE presents a straightforward average of prediction error magnitudes, ensuring a comprehensive grasp of typical discrepancies. As a metric in percentage terms, MAPE quantifies the relative magnitude of forecast errors, indicating the proportional difference between actual and predicted values. Together, these metrics ensure a balanced and thorough appraisal of the model's accuracy and reliability in forecasting cryptocurrency prices.

After evaluating the predictive performance of the forecasting models within the proposed framework, we shift our focus to understanding the underlying price formation factors of Bitcoin and Altcoins. This transition leads us to the final segment of our framework: the SHAP analysis. Through SHAP analysis of the deep learning models, we aim to provide forecasting explainability by exploring the Global and Local feature importance in order to understand the factors responsible for price formation of cryptocurrencies under study.

While the SHAP analysis is embedded in the Algorithm 2, its application and exploration are reserved for Objective 2. By utilizing a consistent framework, we ensure a seamless transition between objectives, focusing on the SHAP analysis's in-depth exploration of the price formation factors of cryptocurrencies discussed in the subsequent section.

3.2.2 Objective 2

02: To determine the price formation factors of Bitcoin and Altcoins

3.2.2.1 Explainable AI Framework to Determine the Price Formation Factors

Objective 2 seeks to identify the factors determining the price of Bitcoin and Altcoins. While the predictive framework detailed in Objective 1 offers a mechanism to forecast prices, it is equally important to understand the underlying determinants of these predictions. Given the complex nature of cryptocurrency markets and the high dimensionality and non-linear characteristics of the input data used for our analysis, traditional statistical methods often fall short of providing clear insights. While deep learning models are competent at dealing with such high-dimensional data, they operate as "black boxes" that conceal the underlying decision-making processes. To bridge this interpretability gap and provide a transparent understanding of feature contributions in the forecasting process, we employ the SHAP (SHapley Additive exPlanations) methodology (Ning et al., 2022). In this section, we discuss the theoretical background of the SHAP methodology and how it addresses the interpretability challenge. SHAP is based on cooperative game theory and assigns a value to each feature by measuring its contribution to the prediction outcome. By using the SHAP DeepExplainer, we can quantify the impact of each feature on the forecasted result, allowing us to gain a deeper understanding of the model's decision-making process.

It is crucial to highlight that the methodologies for predicting the prices of Bitcoin and Altcoins are thoroughly described in Objective 1 of our proposed price forecasting framework, as outlined in Algorithm 2. Objective 1 culminates with model performance evaluation, marking the transition to our subsequent objective. Objective 2 seamlessly extends this framework by incorporating the Shapley analysis. In Objective 2, we investigate the underlying factors shaping the prices of Bitcoin and Altcoins. Figure 6 shows the use of Game Theory-based SHapley Additive exPlanations (SHAP) DeepExplainer approach in the overall framework to identify the features that significantly impact price formation.

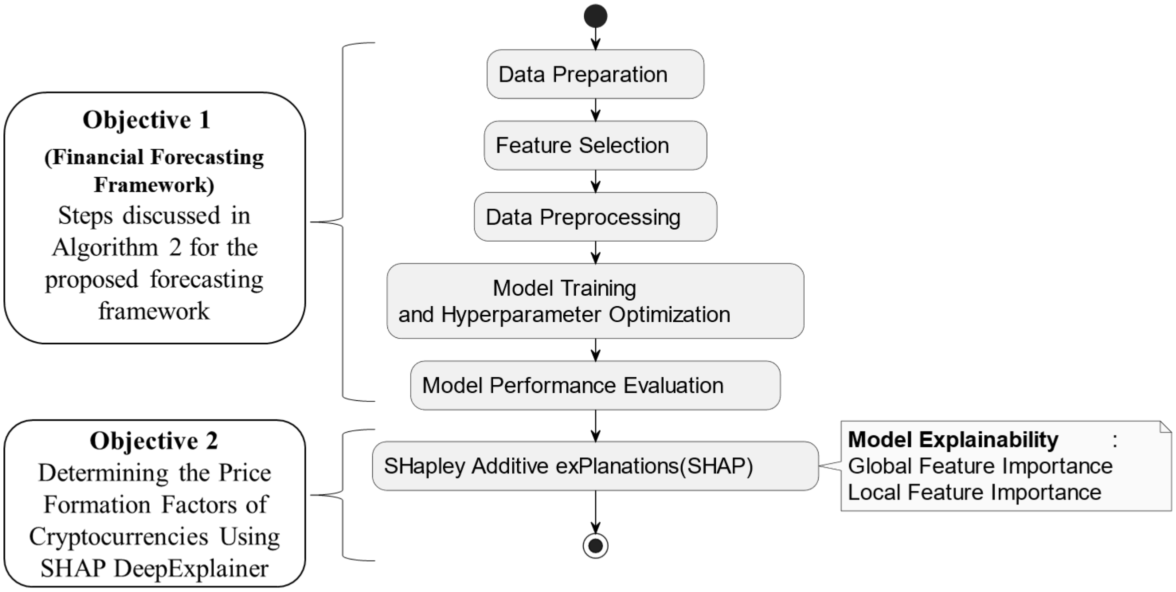


Figure 6. Schematic representation of the analytical framework for determining the price formation factors of Bitcoin and Altcoins

3.2.2.2 DeepSHAP Based Feature Importance

High-dimensional datasets often contain a multitude of potentially relevant features, making it difficult to isolate the true predictors from noise. Traditional statistical methods can struggle with feature selection and multicollinearity in such scenarios, leading to overfitting and reduced model generalizability. Supervised Machine Learning methods face with the problem of overfitting when a high-dimensional input is given. Deep learning models, while capable of handling high-dimensional data, often lack transparency in their decision-making processes. This opacity can obscure the relationships between features and predictions, making it challenging to identify the genuine price formation factors. SHAP (SHapley Additive exPlanations) is a method derived from cooperative game theory to explain individual predictions of machine learning models (Lundberg & Lee, 2017). It computes the contribution of each feature to every prediction, ensuring a fair distribution of the prediction value across features. In cooperative game theory, the Shapley value is a concept that allocates a payout based on the contribution of each player in a coalition. Analogously, in machine learning, SHAP values distribute the prediction value among the features. The Shapley value ensures each feature's allocation is done fairly, based on its contribution. Mathematically, the Shapley value for a feature i, is defined as (Lundberg & Lee, 2017):

$$\phi_i(f) = \sum_{S\subseteq N\setminus\{i\}} \frac{|S|!|(|N|-|S|-1)!}{|N|!}[f(S \cup \{i\}) - f(S)] \tag{45}$$

where, $\phi_i(f)$ represents the Shapley value of feature i. N is the set of all features, and S is a subset of N that does not include feature i. The function f represents the prediction function, and the term $f(S \cup \{i\}) - f(S)$ indicates the marginal contribution of feature S when added to subset S.

DeepLIFT (Deep Learning Important Features) is a method to decompose the output prediction of a neural network on a specific input by backpropagating the contributions of all neurons in the network to every feature of the input (Shrikumar et al., 2017). It answers how much does each input feature change the output. DeepLIFT measures the contribution of each feature by comparing its activation to a reference activation. Unlike the standard backpropagation approach that can sometimes provide misleading importance scores due to non-linearities in the model, DeepLIFT computes the differences in activations between the inputs and a reference input, then backpropagates these differences through the network. This "difference-from-reference" method offers a more intuitive importance score for each feature. DeepLIFT calculates the contribution of each feature by comparing its activation to a reference activation. The contribution C of feature i for a given output o is given by (Shrikumar et al., 2017):

$$C_i^o = \left(x_i - x_i^{ref}\right) \times \Delta F_i^o \tag{46}$$

where x_i is the activation of feature i, x_i^{ref}, is the reference activation for feature i, F_{-i}^{o} is the difference-from-reference in the output when feature i is present versus absent.

DeepSHAP is a method for interpreting the predictions generated by deep learning models, specifically neural networks. DeepSHAP integrates two foundational methodologies: DeepLIFT and Shapley values, aiming to amplify the interpretability of deep learning models. This combination is particularly potent because it amalgamates the power of DeepLIFT in handling deep neural architectures with the fairness guarantees of Shapley values.

DeepSHAP integrates Shapley values with DeepLIFT to ensure a fair distribution of contributions across features. DeepSHAP essentially replaces the difference-from-reference term F_{-i}^{o} in DeepLIFT with the Shapley values. So, the DeepSHAP value for feature i becomes:

$$\phi_i^{\text{DeepSHAP}} = \left(x_i - x_i^{\text{ref}}\right) \times \phi_i\,(f) \tag{47}$$

By combining the difference-from-reference approach of DeepLIFT with the cooperative game theory foundation of Shapley values, DeepSHAP provides robust and interpretable importance scores for features in deep learning models. Implementing DeepSHAP for our analysis on Bitcoin and Altcoins, we generate visualizations that depict both global and local feature importance. These visual representations offer a clear perspective on the relative significance of each feature in determining cryptocurrency prices. The global feature importance provides an overarching view of the consistent influencers across the dataset, while the local importance highlights the role of features for specific data instances. By interpreting these visualizations, we gain a nuanced understanding of the key determinants driving cryptocurrency prices, aiding in both research insights and practical decision-making in the cryptocurrency market.

3.2.3 Objective 3

03: To identify the factors responsible for price inconsistencies across different major cryptocurrency exchanges

3.2.3.1 Identifying Factors for Price Inconsistencies in Cryptocurrency Exchanges

Our third objective is to identify the factors responsible for price inconsistencies across different major cryptocurrency exchanges, specifically for the cryptocurrencies Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), Dash (DASH), and Ripple (XRP). We aim to discern the factors causing price discrepancies for these cryptocurrencies across three major exchanges: Binance, Coinbase, and Kraken. Our datasets for this analysis encompass variables from Table 8, which capture essential price metrics and indicators reflecting the cryptocurrency's online presence and transactional attributes. Additionally, we incorporate all technical indicators listed in Table 1 from Type IDs 3 to 7.

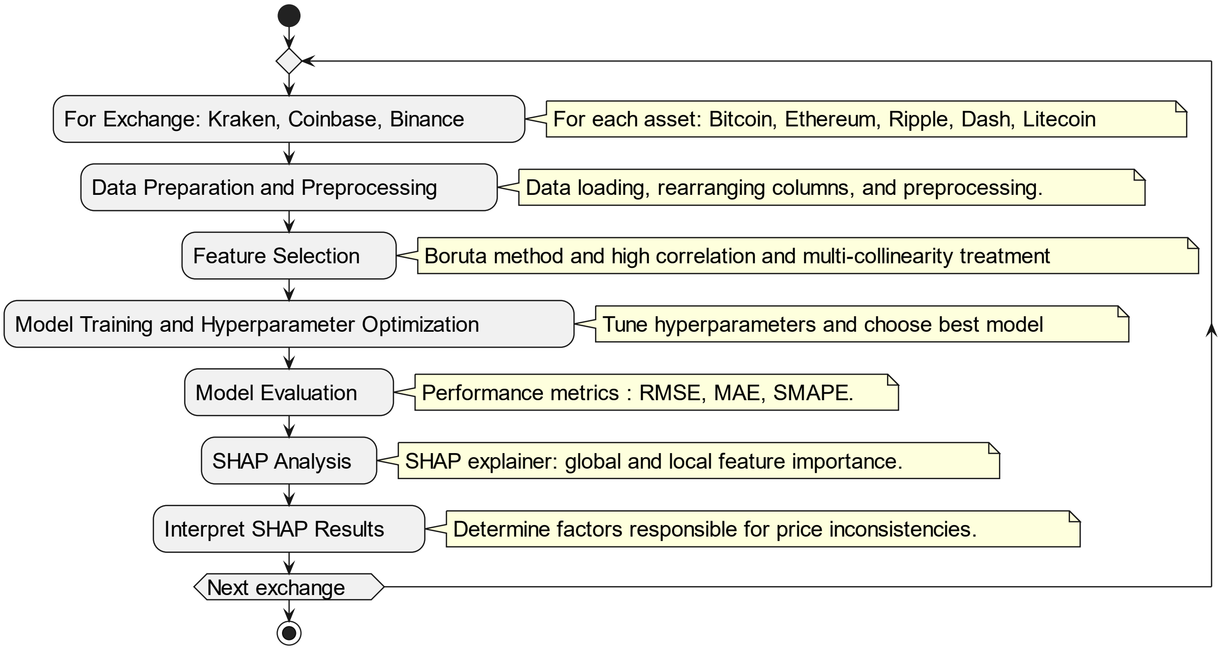


Figure 7. Workflow for Analyzing Cryptocurrency Price Inconsistencies

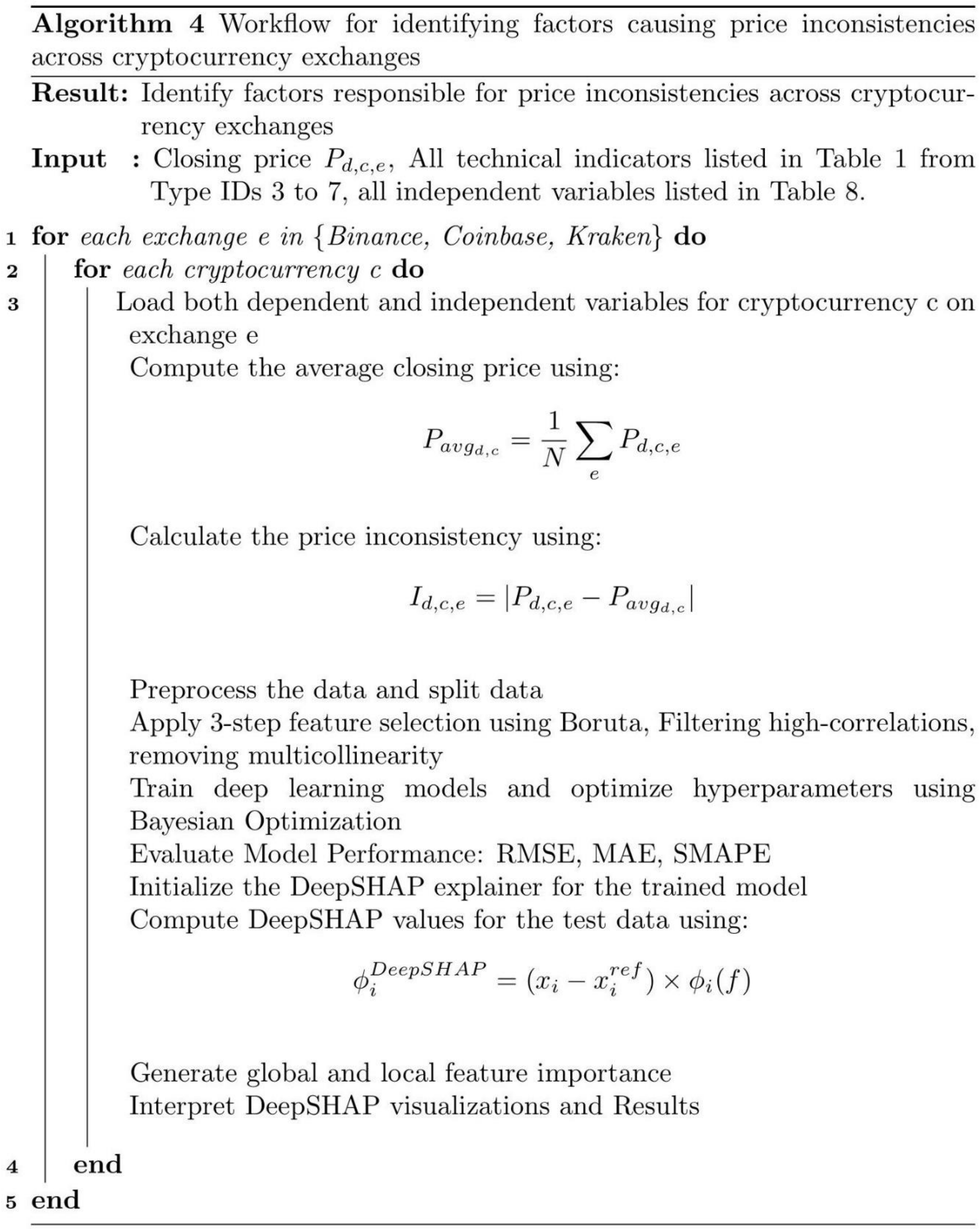
Figure 7 and Algorithm 4 present a self-explanatory systematic workflow of the entire process. This procedure is repeated for each cryptocurrency on each exchange to ensure a comprehensive analysis across all selected exchanges.

The first step in the algorithm involves loading both the dependent and independent variables for the specific cryptocurrency on the selected exchange. Once the data is loaded, the average closing price $P_{avg_{d,c}}$ is computed using the formula $P_{avg_{d,c}} = \frac{1}{N} \sum_{e} P_{d,c,e}$ where the sum is over all exchanges e, and N is the total number of exchanges.

Following this, we calculate the price inconsistency $I_{d,c,e}$ as the absolute difference between the closing price and the average closing price, given by: $I_{d,c,e} = \left| P_{d,c,e} - P_{avg_{d,c}} \right|$.

The data is then preprocessed, and a split is made between training and test datasets. Feature selection is applied in a three-step process: using the Boruta method, filtering out features with high correlations, and removing multicollinearity. The selected features are then used to train deep learning models. We utilize the same hyperparameter search space for Bayesian Optimization as listed in Table 10, akin to the approach we adopted for Objective 2. The performance of these models is then evaluated using metrics such as RMSE, MAE, and SMAPE.

After model evaluation, the DeepSHAP explainer is initialized for the trained model and DeepSHAP values for each feature are computed using the equation (46). The final steps involve generating both global and local feature importance visualizations using DeepSHAP explainer. These visualizations are then interpreted to identify the primary factors driving the observed price inconsistencies across the selected cryptocurrency exchanges. Algorithm 4 below summarizes the entire procedure in a concise and self-explanatory manner.



Algorithm 4 Workflow for identifying factors causing price inconsistencies across cryptocurrency exchanges

Result: Identify factors responsible for price inconsistencies across cryptocurrency exchanges

Input : Closing price $P_{d,c,e}$, All technical indicators listed in Table 1 from Type IDs 3 to 7, all independent variables listed in Table 8.

```
for each exchange e in {Binance, Coinbase, Kraken} do
    for each cryptocurrency c do
        Load both dependent and independent variables for cryptocurrency c on exchange e
        Compute the average closing price using:
            $P_{avg_{d,c}} = \frac{1}{N} \sum_{e} P_{d,c,e}$
        Calculate the price inconsistency using:
            $I_{d,c,e} = |P_{d,c,e} - P_{avg_{d,c}}|$
        Preprocess the data and split data
        Apply 3-step feature selection using Boruta, Filtering high-correlations, removing multicollinearity
        Train deep learning models and optimize hyperparameters using Bayesian Optimization
        Evaluate Model Performance: RMSE, MAE, SMAPE
        Initialize the DeepSHAP explainer for the trained model
        Compute DeepSHAP values for the test data using:
            $\phi_i^{DeepSHAP} = (x_i - x_i^{ref}) \times \phi_i(f)$
        Generate global and local feature importance
        Interpret DeepSHAP visualizations and Results
    end
end
```

where $P_{d,c,e}$ is the closing price for a specific date *d*, for coin *c*, and exchange *e;* $P_{avg_{d,c}}$is the average closing price for date *d*, for coin *c*, and exchange *e;* $I_{d,c,e}$ is the price inconsistency for date *d*, for coin *c*, and exchange *e;*

3.2.4 Objective 4

04: To compare the liquidity of cryptocurrency exchanges with the liquidities of different sizes of selected stocks by using Martin Liquidity Index (MLI)

In this objective, we focus on analyzing and comparing cryptocurrency exchange liquidity with selected stocks of varying sizes, employing the Martin Liquidity Index (MLI) (Loi, 2017) as the baseline metric. The scope of our liquidity investigation focuses primarily on the three prominent volume-traded cryptocurrency exchanges, namely Binance, Coinbase, and Kraken. These liquidities of exchanges are then compared against the liquidities of prominent stock exchanges, namely the NYSE, NASDAQ, NIFTY, and BSE SENSEX. This objective is motivated by a dual purpose. First, we aim to determine the relative liquidities of major cryptocurrency exchanges vis-à-vis traditional stock markets. Second, we aim to identify which among the top cryptocurrency exchanges commands the highest liquidity.

3.2.4.1 Multifaceted Liquidity Assessment

In contrast to traditional stock exchanges, the cryptocurrency market exhibits unique liquidity dynamics. Cryptocurrencies are known for their heightened volatility, leading to rapid and significant price swings within short time frames, affecting the immediacy with which assets can be traded without causing substantial price impact. Unlike traditional stock exchanges with specific trading hours, cryptocurrency exchanges operate round the clock. This continuous trading can lead to non-uniform liquidity patterns, with certain periods experiencing thinner liquidity than others. The regulatory environment for cryptocurrencies is still evolving and varies significantly across jurisdictions, leading to fragmented liquidity as certain exchanges might be inaccessible to investors from specific regions. Furthermore, the cryptocurrency market attracts diverse participants, from retail investors to institutional players, each with different trading behaviors and liquidity needs. Given these unique characteristics, relying solely on a single liquidity measure might not capture the complete picture of the cryptocurrency market's liquidity landscape. To comprehensively assess liquidity in both traditional stock and cryptocurrency markets, we utilize several established liquidity measures from prior research. These measures not only facilitate a comparative assessment of MLI but also enrich our analysis by offering varied perspectives on liquidity, ensuring a more holistic and informed evaluation.

In this objective, we use the Martin Liquidity Index (MLI) (Loi, 2017) as our baseline measure of liquidity. The MLI examines the relationship between price changes and trading volume, indicating how price variations can affect trading volume and, in turn, market liquidity. Alongside the MLI, we employ the Amihud Illiquidity Ratio (AIR), a recognized metric in finance that measures stock illiquidity by comparing the absolute return of a stock to its trading volume (Amihud, 2002). The AR Bid-Ask Spread, developed by (Abdi & Ranaldo, 2017), provides an estimate of the bid-ask spread using available price data. The spread estimator by

(Corwin & Schultz, 2012) uses daily price highs and lows to estimate bid-ask spreads. Additionally, we use the Richard Rolls measure (ROLL, 1984), which calculates the effective bid-ask spread from observed market price changes in efficient markets. These methods, each with its own approach and data requirements, together offer a comprehensive view of liquidity to address our research questions. Figure 8 shows the method we use to calculate liquidity using the Martin Liquidity Index (MLI) for each exchange, and a similar approach is followed for the other four liquidity measures.

3.2.4.2 Proposed Framework for Liquidity Analysis

3.2.4.2.1 Experimental Settings

For our experiments, we utilize the trading data from three major cryptocurrency exchanges, focusing on the top volume-traded currency pairs of each exchange. It is essential to clarify that our dataset includes a mix of both crypto-to-fiat and crypto-to-crypto currency pairs. Stablecoins are also part of our analysis. Pairs like BTCUSD and BTCEUR are treated as separate pairs, given that they represent different trading behaviors and liquidity dynamics in their respective fiat markets. On Kraken, we analyze the top 150 volume-traded currency pairs, accounting for 98.48% of the total traded volume. For Coinbase, our dataset encompasses the top 167 volume-traded currency pairs, representing 99.02% of the total traded volume. On Binance, we examine the top 200 currency pairs, making up 94.63% of the traded volume. It is important to note that our primary data source for Binance is sourced from binance.com, the main global exchange, excluding dedicated exchanges for specific regions such as binance.co.uk and binance.us. By concentrating on the top volume-traded currency pairs, we effectively capture the most stable and consistently traded assets and exclude new, unstable, and speculative coins that might be subject to pump-and-dump schemes or other volatile trading behaviors. These speculative currency pairs often lack long-term reliability in liquidity measurement, especially as some may cease trading due to the associated startup or company's discontinuation. While they might exhibit brief periods of high performance, their overall consistency is questionable. Their exclusion ensures that our analysis remains robust, focusing on established and stable assets that reflect genuine trading behaviors and liquidity patterns of the exchange.

The data is extracted, cleaned, and normalized using Min-Max normalization for each currency pair. The data is further segmented and analyzed over three distinct intervals, each representing a unique market phase, as detailed in Table 9 (listed under Section 3.1.3 Dataset for Objective 4). These intervals have been chosen to capture the evolving dynamics of the cryptocurrency market over time, ensuring a comprehensive understanding of liquidity across different market conditions. A detailed list of each currency pair used in the study is provided in Section 3.1.3 Dataset.

3.2.4.2.2 Liquidity Calculation Methods

i) Martin Liquidity Index (MLI)

We use Martin Liquidity Index (MLI) as the baseline metric for liquidity calculation. Figure 8 shows the overall step-by-step workflow of the liquidity calculation process. We begin by loading each currency pair for a given exchange under consideration for the given time interval. Upon loading the data, the Martin Liquidity Index (MLI) for each currency pair is calculated in accordance with below equation:

$$MLI_t = \sum_{i=1}^{N} \frac{(P_{it} - P_{i(t-1)})^2}{V_{it}} \tag{48}$$

where P_{it} is the closing price of the ith asset at time t and V_{it} is the negotiated volume of the trades for the ith asset at time t and N is the total number of assets traded on the exchange.

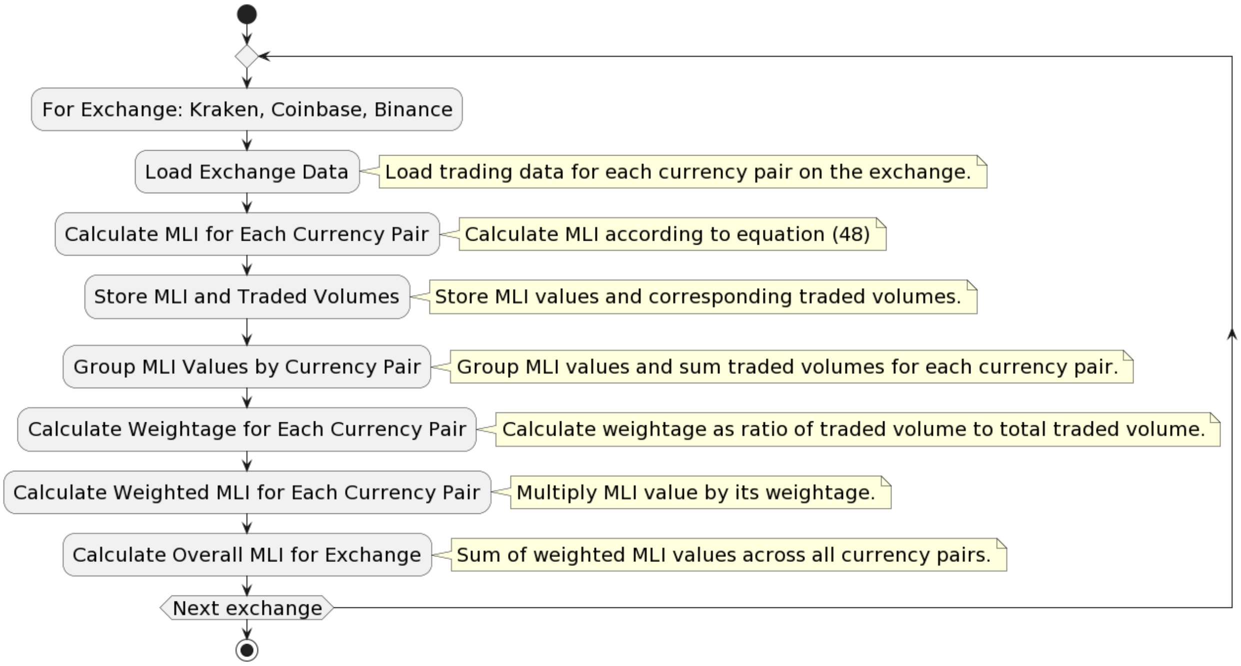


Figure 8. Workflow for calculating Martin Liquidity Index (MLI) of different cryptocurrency exchanges

We then store the MLI values along with the corresponding traded volumes. We group these values by currency pair and sum the traded volumes for each pair. Next, we calculate the weightage of each currency pair as the ratio of its traded volume to the total traded volume across all pairs. We then use this weightage to calculate the weighted MLI for each pair, which is the product of its MLI value and its weightage. Finally, we calculate the overall MLI for the exchange by taking the sum of the weighted MLI values across all currency pairs. This overall MLI provides a measure of the exchange's liquidity, considering each currency pair's trading volume. Mathematically, the computation of the overall MLI value can be expressed as follows:

$$Weighted\ MLI_i = MLI_i \times \frac{Traded\ Volume\ _i}{Total\ Traded\ Volume} \tag{49}$$

$$Overall\ MLI\ = \sum_{i=1}^{N}\ Weighted\ \ MLI\ _i \tag{50}$$

where $Weighted\ MLI_i$ is the weighted Martin Liquidity Index for the i^{th} currency pair, $MLI\ _i$ is the Martin Liquidity Index for the i^{th} currency pair, $Traded\ Volume\ _i$ is the traded volume for the i^{th} currency pair, Total Traded Volume is the sum of traded volumes across all currency pairs.

The incorporation of traded volume as a weighting factor in the MLI computation is grounded in the fundamental principle of market liquidity. It is widely recognized that assets with higher trading volumes naturally exhibit higher levels of liquidity. In the context of an exchange, each currency pair's trading volume serves as a direct indicator of its liquidity and, by extension, its influence on the overall liquidity of the exchange. When we consider the MLI, which quantifies the relationship between price changes and trading volume, it becomes evident that not all currency pairs contribute equally to the exchange's liquidity profile. By giving weights to the MLI values based on traded volume, we effectively attribute a proportional level of importance to each currency pair, considering its contribution to overall liquidity. This ensures that currency pairs with limited trading volume do not skew our overall liquidity assessment but accurately reflect the exchange's liquidity dynamics. We repeat the above process for each cryptocurrency exchange (Kraken, Coinbase, Binance) and traditional stock exchanges (NYSE, NASDAQ, BSE SENSEX and NIFTY) under the study. Figure 8 shows the overall step-by-step workflow of the liquidity calculation process.

A high MLI value indicates a greater price dispersion relative to the quantity traded, which can affect trading volume and, in turn, market liquidity. A high MLI value means lower liquidity, and a low MLI value suggests higher liquidity, meaning price variations have a lesser effect on trading volume.

ii) Amihud's Illiquidity Ratio

Amihud's Illiquidity Ratio (AIR), a widely used measure of price impact, is computed as the annual mean of absolute daily returns relative to daily dollar trading volume. In contrast to the bid-ask spread, this index possesses the advantage of being more readily computable due to the broad accessibility of requisite data, even in markets lacking sophisticated spread measures. The Amihud Illiquidity Ratio to measure the market breadth is given as (Amihud, 2002):

$$AIR_T^i = \frac{1}{D_T} \sum_{t=1}^{D_T} \frac{|R_{t,T}^i|}{Vol_{t,T}^i} \tag{51}$$

where D_T is the number of days for which data are available, $R_{t,T}^{i}$ is the return on day t of the year T, and $Vol_{t,T}^{i}$ is the daily volume. An elevated Amihud Illiquidity Ratio (AIR) indicates a narrow market scope and limited liquidity, while a lower AIR indicates a broad market scope and substantial liquidity.

To determine the liquidity at the exchange level, we employ a similar volume-weighted liquidity computation procedure for Amihud's Illiquidity Ratio (AIR) as was applied to the MLI. The overall AIR at exchange level can be expressed as:

$$Overall\ AIR = \sum_{i=1}^{N} Weighted\ \ AIR_{\ i} \tag{52}$$

where N is the total number of currency pairs or assets traded on the exchange

A high AIR value signifies lower liquidity, indicating larger price changes relative to the trading volume. In contrast, a low AIR value signifies higher liquidity, suggesting minimal price changes given the trading volume.

iii) AR Bid-Ask Spread

(Abdi & Ranaldo, 2017) developed a method to estimate bid-ask spreads using the daily high, low, and closing prices. This technique is particularly useful when direct quote data is unavailable. The method provides a reliable estimation of the bid-ask spread by focusing on these three-price metrics. Importantly, it is well-suited for cryptocurrency markets due to the widespread availability of OHLCV (Open, High, Low, Close, Volume) data for each cryptocurrency exchange under consideration. The mathematical foundation of this method ensures precision, and its design allows for straightforward computation.

The key aspect of their approach centers on identifying the midpoint between the logarithmic high and low prices within a given time frame. This midpoint is represented mathematically as:

$$\bar{p}_t = \frac{h_t + l_t}{2} \tag{53}$$

where $h_t = \ln(H_t)$ is the logarithm of the high price, $l_t = \ln(L_t)$ represents the logarithm of the low price, and $c_t = \ln(C_t)$ denotes the logarithm of the closing price.

The authors suggested two spread calculation variations: a monthly and a two-day corrected version. In our analysis, we employ the two-day adjusted version. This variant includes both high and low prices from two consecutive intervals, t and $t + 1$. The estimator, denoted as AR_i, is calculated as follows (Abdi & Ranaldo, 2017):

$$AR_i = \sqrt{max\{4(c_t - \bar{p}_t)(c_t - \bar{p}_{t+1}),0\}} \tag{54}$$

The calculation of the volume-weighted average for each currency pair can be represented as:

$$\text{Weighted } AR_i = AR_i \times \frac{\text{Total Volume for the period}}{\text{Daily Volume}_i} \tag{55}$$

The weighted bid-ask spreads of all currency pairs provide a consolidated measure of the exchange's overall liquidity. The exchange-level liquidity, AR_{exchange} can be expressed as follows:

$$AR_{\text{exchange}} = \sum_{j=1}^{N} \text{Weighted } AR_{i,j} \tag{56}$$

A high AR Bid-Ask Spread value indicates a lower liquidity. On the other hand, a low AR Bid-Ask Spread denotes better liquidity, as the prices of buying and selling are more closely matched.

iv) Roll's Covariance Liquidity Estimator

(ROLL, 1984) introduced a liquidity metric designed to estimate the bid-ask spread using observable sequential price changes indirectly. The foundation of this method lies in the understanding that price shifts are influenced by two main factors: the hidden bid-ask spread and price adjustments triggered by new information. In scenarios where positive and negative informational events have an equal likelihood, their impact on price changes tends to neutralize. As a result, the bid-ask spread becomes discernible by examining the covariance of consecutive price changes.

Mathematically, Roll's measure, Roll_{it}, for each asset i, can be expressed as follows:

$$\text{Roll}_{it} = \begin{cases} 2\sqrt{-\text{Cov}\,(\Delta p_{it}; \Delta p_{it-1})} & \text{if Cov}\,(\Delta p_{it}; \Delta p_{it-1}) < 0 \\ 0 & \text{if Cov}\,(\Delta p_{it}; \Delta p_{it-1}) \geq 0 \end{cases} \tag{57}$$

This method proves useful when the bid-ask spread is neither directly observable nor easily accessible. By analyzing the covariance between successive price changes, investors can derive insights into the concealed bid-ask spread, enabling more informed trading decisions. Moreover, this approach offers a dynamic perspective on the bid-ask spread, adapting as market conditions and the influx of information evolve.

To determine the liquidity at the exchange level, we employ a similar volume-weighted liquidity computation procedure for Roll's measure as was applied to the preceding methods. $Roll_{\text{exchange}}$ at the exchange level can be expressed as the summation of all volume weighted liquidities for each currency pair i:

$$Roll_{\text{exchange}} = \sum_{i=1}^{N} \text{ Weighted Roll}_i \tag{58}$$

A high Roll's Measure value indicates that market price fluctuations are considerably greater than the effective bid-ask spread, indicating a lower level of liquidity. On the other hand, a low Roll's Measure value indicates increased liquidity.

v). CS spread estimator

The (Corwin & Schultz, 2012) estimator calculates the bid-ask spread using daily high and low prices. The underlying premise of this method is that daily high prices typically result from buyer-initiated trades, while seller-initiated trades influence low prices. Consequently, the day's high-to-low price ratio encompasses both the intrinsic volatility of the security and its bid-ask spread. By analyzing this ratio, the estimator can distinguish the spread's influence from the inherent price volatility of the security. The core principle hinges on the observation that while the intrinsic volatility scales with the period under study, the spread component remains relatively stable. This distinction facilitates the simultaneous estimation of both the spread and the volatility.

The methodology employs two equations: one derived from the high-low ratios of two consecutive days and another from the high-low ratio over a two-day span. The estimator can efficiently distinguish the impact of the spread from the security's inherent price volatility by examining this ratio. The spread component remains largely unchanged, although the inherent volatility grows with the duration considered in this methodology. This distinctive behavior enables the concurrent calculation of the dispersion and the volatility. The following equations can describe the estimator (Corwin & Schultz, 2012) :

Beta (β) captures the squared logarithmic ratios of high to low prices over two consecutive days:

$$\beta = \left[\ln\left(\frac{H_i}{L_i}\right)\right]^2 + \left[\ln\left(\frac{H_{i+1}}{L_{i+1}}\right)\right]^2 \tag{59}$$

where H_i and L_i are the high and low prices on day i.

Gamma (γ) represents the squared logarithmic ratio of the highest to the lowest prices over two days:

$$\gamma = \left[\ln\left(\frac{H_{i,i+1}}{L_{i,i+1}}\right)\right]^2 \tag{60}$$

where $H_{i,i+1}$ and $L_{i,i+1}$ are the highest and lowest prices over days i and $i + 1$.

From β and γ, Alpha (α) is calculated. α serves as crucial intermediate step in the methodology, and its value is used to compute the final estimator, $CS_{i,i+1}$, which represents the bid-ask spread. α is given by:

$$\alpha = \frac{\sqrt{2\beta}-\sqrt{\beta}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma}{3-2\sqrt{2}}} \tag{61}$$

where α is a derived intermediate parameter that effectively integrates the price dynamics over two days into a single parameter by capturing the combined effects of the bid-ask spread and the stock's volatility on the observed high and low prices over two consecutive days.

The final bid-ask spread estimator, $S_{i,i+1}$, is computed as:

$$CS_{i,i+1} = \frac{2(\exp(\alpha)-1)}{1+\exp(\alpha)} \tag{62}$$

Next, we compute the volume-weighted average liquidity for each currency pair as follows:

$$\text{Weighted } CS_{i,i+1} = CS_{i,i+1} \times \frac{\text{Total Volume for the period}}{\text{Daily Volume}_{i,i+1}} \tag{63}$$

The overall liquidity at the exchange level, $CS_{exchange}$, is calculated as the weighted sum of:

$$CS_{exchange} = \sum_{j=1}^{N} \text{Weighted } CS_{i,i+1,j} \tag{64}$$

A high value of the CS Spread measure denotes an increased bid-ask spread, indicative of decreased liquidity. A low value suggests a narrower bid-ask spread, indicating higher liquidity.

Algorithm 5 presents a complete methodology for computing different liquidity measures for exchanges. The process commences by sequentially iterating through each exchange, followed by iterating through each asset within each exchange. The Martin Liquidity Index (MLI) is computed for every currency pair, serving as the primary metric. Subsequently, a comparative analysis is conducted with Amihud's Illiquidity Ratio (AIR), AR Bid-Ask Spread, Roll's Covariance Liquidity Estimator, and the CS spread estimator. This methodology offers a thorough evaluation of liquidity to address the research questions for this objective.

Algorithm 5 Computation of MLI, AIR, AR, Roll, CS, and their weighted and overall values for Exchanges

Result: Compute MLI, AIR, AR, Roll, CS liquidity values for Exchanges
Input : Closing price P_{it}, negotiated volume V_{it}, returns $R^i_{t,T}$, daily volume $Vol^i_{t,T}$, and mid price $\bar{p}_t$ for all i^{th} assets, for all exchanges

```
for each exchange do
    for each asset on the exchange do
        Compute $MLI_t$ as $MLI_t = \sum_{i=1}^{N} \frac{(P_{it}-P_{it-1})^2}{V_{it}}$
          Compute Weighted $MLI_i$ as $MLI_i \times \frac{V_{it}}{\sum_{j=1}^{N} V_{jt}}$
        Compute $AIR^i_T$ as $AIR^i_T = \frac{1}{D_T} \sum_{t=1}^{D_T} \frac{|R^i_{t,T}|}{Vol^i_{t,T}}$
          Compute Weighted $AIR_i$ as $AIR_i \times \frac{V_{it}}{\sum_{j=1}^{N} V_{jt}}$
        Compute $AR_i$ as $AR_i = \sqrt{\max\{4(c_t - \bar{p}_t)(c_t - \bar{p}_{t+1}), 0\}}$
          Compute Weighted $AR_i$ as $AR_i \times \frac{TotalVolume}{DailyVolume_i}$
        Compute $Roll_{it}$ based on covariance condition
          Compute Weighted $Roll_i$ as $Roll_i \times \frac{V_{it}}{\sum_{j=1}^{N} V_{jt}}$
        Compute $CS_{i,i+1}$ using the given formula
          Compute Weighted $CS_{i,i+1}$ as $CS_{i,i+1} \times \frac{TotalVolume}{DailyVolume_{i,i+1}}$
    end
    Overall Liquidity Calculations:
      Compute overall MLI for the exchange using eq (50)
      Compute overall AIR for the exchange using eq (52)
      Compute overall AR for the exchange using eq (56)
      Compute overall Roll for the exchange using eq (58)
      Compute overall CS for the exchange using eq (64)
end
```

Chapter 4

Results and Discussion

This chapter presents the detailed findings and analyses corresponding to each research objective outlined in previous chapters. The results are presented, interpreted, and compared to baseline models and relevant literature. The goal is to offer a thorough understanding of the research outcomes, focusing on cryptocurrency trends, price formation factors, inconsistencies across exchanges, and liquidity comparisons.

4.1 Objective 1: Trends and Forecasts of Bitcoin and Altcoin Prices

4.1.1 Results for Objective 1

01: To analyze the trends of exchange prices of Bitcoin and Altcoins and to forecast future exchange prices of Bitcoin and Altcoins

In this section, we first present the results for analyzing the trends of exchange prices for Bitcoins and Altcoins and discuss their interpretation. Next, we present the results obtained from forecasting framework illustrated in Figure 1.

4.1.1.1 Results for Cryptocurrency Price Trends

This section details the results of Cryptocurrency price trends for each of the four coins using Change point Analysis with the PELT algorithm. It also presents the findings from the Multifractal Detrended Fluctuation Analysis.

4.1.1.1.1 Results of Change point Analysis (PELT)

Change Point are significant junctures in a time series where the statistical properties of the data change. In Tables 11-15, each change point demarcates a distinct period or segment in the asset's price trend. Each segment is characterized by its own mean, standard deviation, kurtosis, and skewness, providing a unique snapshot of asset's price behavior during that period. The tables below present a segmented analysis of closing prices for Bitcoin and Altcoins. Each row corresponds to a distinct period, demarcated by significant changes in the price trend.

Table 11. Descriptive statistics for Change Points obtained from PELT algorithm for Bitcoin (BTC) from 2-Jan-17 to 2-Apr-23.

Changepoint	Mean	Median	Std Dev	Kurtosis	Skewness
1	1680.2871	1253.0800	728.6321	-1.2666	0.5302
2	5211.9595	4578.8950	1410.6134	-0.5509	0.7941
3	13981.1879	14026.5996	2608.3745	-0.9544	0.1612
4	7581.4026	7042.3700	1334.0772	0.2184	1.0267

5	4232.1588	3945.2519	788.8663	0.9254	1.2606
6	9324.3881	9318.4908	1620.1929	-0.2607	0.0449
7	18470.0130	18370.0025	2793.3743	-0.3322	0.4590
8	34412.5831	34269.5215	4140.5519	0.1586	0.1052
9	54124.3626	55792.0323	4913.8368	-0.8836	-0.3420
10	35868.4165	35584.1932	2989.3048	-0.7641	0.2606
11	46651.0930	47050.9947	2634.6409	-0.3543	-0.2061
12	59968.6530	60657.2006	3559.2894	-0.8359	0.0112
13	48496.1875	48157.6607	1627.2655	-1.5188	0.1696
14	41019.7645	41077.9980	2879.5646	-0.6526	0.2153
15	30015.4354	29832.9142	1109.0495	2.9473	1.4266
16	19723.9537	19796.8094	2272.1593	-0.6360	0.2082
17	24134.0052	23475.4667	2200.6358	-0.5653	0.6846

Source: Self Compilation.

Table 12. Descriptive statistics for Change Points obtained from PELT algorithm for Ethereum (ETH) from 2-Jan-17 to 2-Apr-23

Changepoint	Mean	Median	Std Dev	Kurtosis	Skewness
1	36.6047	22.4551	29.3131	0.1737	1.0041
2	300.5755	297.4750	70.6703	0.3694	0.4706
3	910.4446	867.6925	178.7919	-0.0133	0.6496
4	541.5977	519.3165	109.5526	-0.7230	0.5382
5	191.2730	185.9128	53.8179	-0.2398	0.4228
6	459.8831	401.8663	125.1882	3.3186	1.6449
7	1567.0348	1614.2278	251.5756	-1.1705	-0.3601
8	2479.3211	2349.3318	529.4631	1.4072	1.3575
9	3329.8156	3297.2257	280.1607	-0.2743	0.3663
10	4189.9529	4145.2228	290.5420	-0.8629	0.1960
11	2951.7587	2946.2571	297.9736	-0.3668	0.3018
12	1938.3669	1944.8278	187.1245	1.4489	0.7961
13	1447.9896	1465.7035	220.9413	-1.0949	0.0891

Source: Self Compilation.

Table 13. Descriptive statistics for Change Points obtained from PELT algorithm for DASH from 2-Jan-17 to 2-Apr-23

Changepoint	Mean	Median	Std Dev	Kurtosis	Skewness
1	21.8287	16.7744	12.4998	3.1169	1.8602
2	87.7434	88.1794	17.5421	0.5794	0.7130
3	179.7441	182.9985	22.5348	-0.3301	-0.1557
4	312.4964	309.5100	30.2755	0.1555	0.4930
5	868.3915	804.8865	270.9580	-0.7219	0.1029
6	625.0652	620.4850	59.1493	0.7394	-0.4305
7	398.3675	402.0810	72.7095	-1.0425	0.1930
8	246.4833	243.5370	19.0118	1.4494	0.8322
9	171.8377	168.5180	22.5122	-0.6371	0.0487
10	81.0885	81.7362	10.3426	-0.2909	0.2139
11	135.8282	132.0224	22.2406	-1.3531	0.1719

12	76.9031	73.0410	17.6715	0.9610	0.6662
13	108.6484	104.9112	16.0039	2.7258	1.4908
14	234.1716	226.4979	28.8634	0.7036	1.0268
15	326.2851	316.8456	49.7601	-0.6068	0.5319
16	156.7745	156.1464	26.6146	-0.6674	0.3783
17	199.7563	194.6039	25.1159	0.2497	0.7074
18	114.1806	109.1928	19.2445	-1.3102	0.2423
19	49.9695	47.4716	9.4061	0.2735	0.8507

Source: Self Compilation.

Table 14. Descriptive statistics for Change Points obtained from PELT algorithm for Litecoin (LTC) from 2-Jan-17 to 2-Apr-23

Changepoint	Mean	Median	Std Dev	Kurtosis	Skewness
1	6.4029	4.0404	4.4592	4.1587	2.0381
2	38.8935	42.1360	8.6747	-1.2910	-0.3084
3	66.1153	60.5307	19.3132	5.9055	2.1821
4	271.4631	264.9320	35.6054	-0.3726	0.5583
5	183.3383	180.1400	26.9556	-0.9082	0.0715
6	133.4368	131.8960	17.1089	-0.5958	0.4914
7	84.1145	83.1562	7.0720	-0.0991	0.8096
8	56.2893	56.2740	4.4615	0.0785	0.3327
9	35.8898	32.8271	7.8502	-0.1549	0.8026
10	74.6769	74.9828	12.1804	-1.1303	0.0401
11	121.3137	119.7049	12.4040	-0.6673	-0.2438
12	80.5884	76.2699	11.9996	-1.1832	0.2587
13	51.1718	48.0598	9.4869	0.3875	0.8131
14	87.7134	82.9115	14.4923	0.2617	1.0806
15	144.5065	142.4308	14.4108	-0.6670	0.5648
16	198.8546	197.3654	18.5911	-0.7858	0.3779
17	287.6381	278.2910	42.0229	-0.3979	0.6274
18	176.3714	174.8341	13.4742	0.5929	0.2903
19	134.0495	134.5281	10.4732	-0.1715	-0.1863
20	175.4241	177.6032	15.6466	2.2010	0.5968
21	212.9565	204.4229	24.2648	0.3583	1.1955
22	148.8868	148.8014	9.5916	-0.3904	-0.4647
23	113.3565	110.8571	10.8184	-0.3944	0.6366
24	57.9704	56.6023	6.7427	3.5815	1.2462
25	83.9304	85.5023	10.2068	-1.0097	-0.1938

Source: Self Compilation.

Table 15. Descriptive statistics for Change Points obtained from PELT algorithm for Ripple (XRP) from 2-Jan-17 to 2-Apr-23

Changepoint	Mean	Median	Std Dev	Kurtosis	Skewness
1	0.0162	0.0066	0.0170	3.8290	1.9143
2	0.2258	0.2152	0.0464	0.5276	0.6909
3	1.5763	1.3687	0.7511	-0.1585	0.6552
4	0.9662	0.9544	0.1262	-0.6532	-0.0475

5	0.6837	0.6743	0.1188	-0.7267	0.2420
6	0.4731	0.4614	0.0353	-0.2068	0.8318
7	0.3160	0.3219	0.0358	2.3228	1.0175
8	0.4827	0.4744	0.0465	-0.2771	0.1859
9	0.3444	0.3241	0.0451	0.0033	1.0118
10	0.2363	0.2394	0.0400	-0.8117	-0.0167
11	0.5730	0.5779	0.0583	-0.2094	-0.3815
12	0.2843	0.2788	0.0563	6.4599	2.1861
13	0.4945	0.4736	0.0605	-0.7029	0.2557
14	1.3836	1.4010	0.2452	0.5099	-0.7413
15	0.9363	0.9133	0.0879	0.1465	0.6953
16	0.6747	0.6516	0.0873	2.3090	1.1873
17	1.1025	1.0935	0.0988	-0.0599	0.0904
18	0.7643	0.7747	0.0869	-0.3170	-0.3116
19	0.3914	0.3817	0.0511	0.3973	0.9175

Source: Self Compilation.

The segmented analysis of the Tables 11-15, for closing prices of Bitcoin, Ethereum, DASH, LTC, and XRP reveal several key insights about the dynamics of these cryptocurrencies' price trends:

1. **Bitcoin**: Table 11 shows periods of significant price growth, such as the transition from Change Point 7 (mean price approximately 18470.01) to Change Point 8 (mean price approximately 34412.58). However, it also shows periods of price decline, such as the transition from Change Point 9 (mean price approximately 54124.36) to Change Point 10 (mean price approximately 35868.42). The highest kurtosis is observed at Change Point 14 (kurtosis approximately 2.95), indicating a period of increased likelihood of extreme price changes. Figure 9(a) illustrates a detailed examination of the Bitcoin price trends, as marked by the observed Change Points. It can be observed from both Table 11 and Figure 9(a) that Change Point 6 (skewness: 0.0449) and Change Point 12 (skewness: 0.0112) have skewness values close to zero, suggesting a more symmetrical distribution around the mean, which is indicative of a linear trend. On the other hand, Change Points with significant skewness and kurtosis values are more likely to exhibit non-linear trends. Specifically, Change Point 5 with a skewness of 1.2606 and a kurtosis of 0.9254, Change Point 4 with a skewness of 1.0267 and a kurtosis of 0.2184, and Change Point 15 with a skewness of 1.4266 and a pronounced kurtosis of 2.9473, all suggest non-linear trends in the data. These non-linear trends might be characterized by sudden spikes, troughs, or curvature in the data, as opposed to the steady progression observed in linear trends. The observations from Table 11 and Figure 9(a) corroborate with the findings of (Rashid & Ismail, 2023), who also identified that the Bitcoin price time series exhibits not just nonlinear trend patterns but also linear trends.

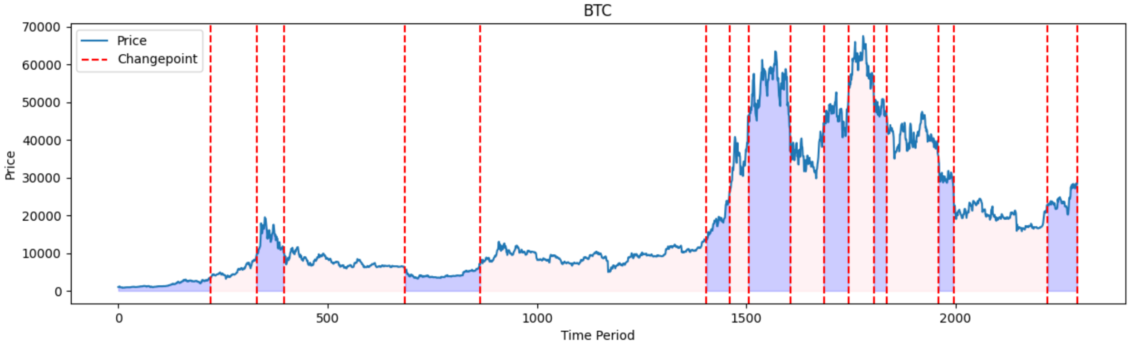


Figure 9 (a). Change Point analysis of the exchange prices of Bitcoin. The plot shows the exchange prices of Bitcoin (BTC) over time with vertical red lines indicating the detected Change Points. The shaded regions between Change Points highlight periods of relative stability in the exchange prices.

Additionally, we conduct the Mann-Whitney U-test to determine whether the observed price changes between these change points hold statistical significance or are merely the result of random fluctuations. Tables A6 to A10 in Appendix A, illustrate the Mann-Whitney U test results for Bitcoin, Ethereum, DASH, Litecoin, and Ripple.

It can be observed from Table A6 that some segments such as 18673.0, 51850.0, and 8000.0 exhibit High U-Statistic values which indicates significant overlaps in price distributions between the compared segments. Economically, such overlaps suggest periods where the Bitcoin market experienced gradual price movements rather than abrupt changes in the initial phase. These phases represent steady market conditions, where external factors did not induce sharp price fluctuations during the initial market phases. The high U-Statistic values, despite the statistical significance indicated by low P-Values, point towards a market characterized by relative stability or consistent trends during these intervals. In contrast, low U-Statistic values (e.g., 0.0 and values close to 0.0) reflect minimal overlap in price distributions, indicating distinct shifts in Bitcoin prices between the intervals. These results signal significant market events or changes, where Bitcoin prices moved distinctly from one period to the next. Such shifts are associated with major market developments, substantial investor reactions to new information, or changes in market sentiment. The low U-Statistic values, along with low P-Values, highlight periods of notable market volatility or significant changes in price trends. These results reveal a mixture of gradual and extreme price changes across different phases, reflecting the complex nature of the Bitcoin market.

2. **Ethereum**:

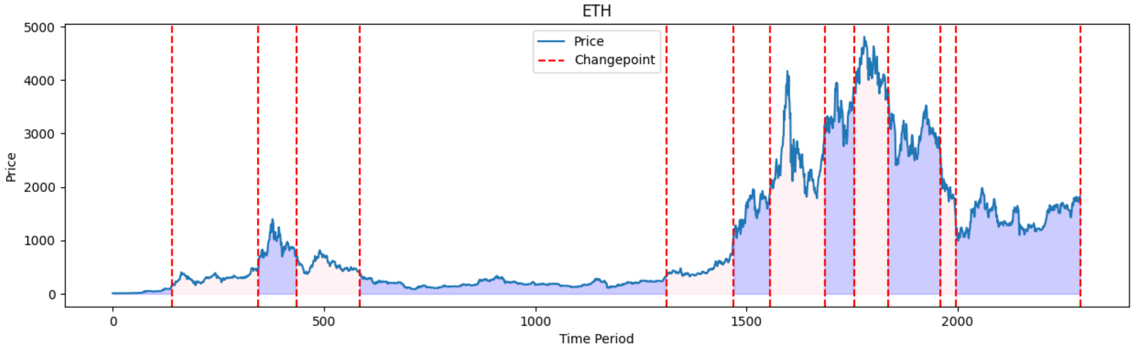


Figure 9 (b). Analytical depiction of changepoints in the Ethereum time series.

Ethereum's price trends also show non-linear behavior. For instance, the transition from Change Point 6 (mean price approximately 625.07) to Change Point 7 (mean price approximately 1567.03) shows a significant increase in the mean price. According to Figure 9(b) and Table 12, the highest kurtosis is observed at Change Point 3 (kurtosis approximately 5.91), indicating a period of increased market risk.

In Table A7 of Appendix A, it is observed that across all the change points, the P-values observed are extremely low, indicating statistical significance even where higher U-Statistic values signal distributional overlap. This complements the economic analysis suggesting notable market shifts between many of the examined intervals. Additionally, the volatility patterns seen in Ethereum echo similar behaviors noted earlier for Bitcoin, indicating that Ethereum price trends follow and are influenced by the movements of the broader cryptocurrency markets and market leader Bitcoin. The stability of intervals such as 435 to 585 for Ethereum align with previously identified periods of gradual pricing for Bitcoin, suggesting ETH tends to shift alongside Bitcoin. At the same time, intervals such as 140 to 345 for Ethereum indicate spikes in volatility also experienced in Bitcoin's history. This implies Ethereum is subject to similar external stimuli and sentiment changes as the pioneer cryptocurrency. Overall, the results indicate that Ethereum's distinctive evolutionary path also exhibits volatility patterns that mirror Bitcoin to some extent, with economic and statistical indicators highlighting the asset's interdependence on the broader cryptocurrency market movements. The significance of all identified P-values underscores the relevance of the change points showing distributional divergences tied to impactful market events.

3. **DASH**: DASH's price trends show significant fluctuations. For instance, the transition from Change Point 5 (mean price approximately 868.39) to Change Point 6 (mean price approximately 625.07) shows a significant decrease in the mean price. The highest kurtosis is observed at Change Point 3 (kurtosis approximately 5.91), indicating a period of increased likelihood of extreme price changes.

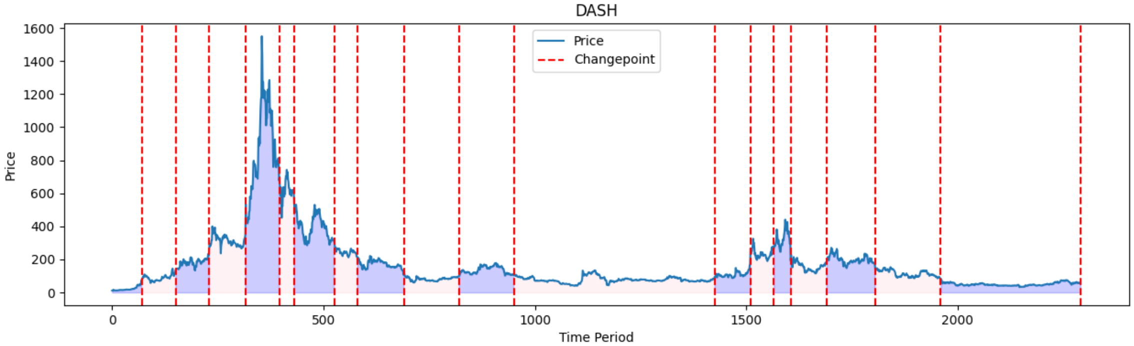


Figure 9 (c). Analytical depiction of Change Points in the DASH time series.

Table A8 in Appendix A, shows the Mann-Whitney U test results for DASH coin. The Dash market evolution shows both steady intervals and periods of extreme distributional divergences like Bitcoin and Ethereum. High U-Statistics for the 820 to 950, 950 to 1425, and 1805 to 1960 change points indicate pricing stability likely tied to consistent external conditions. Meanwhile, extremely low U-Statistics near 0 between several other periods (e.g. 150 to 230, 315 to 395, 1510 to 1565) point to volatility from major events. The uniformly significant P-Values confirm the wider market shifts. Initial surges and fluctuations settle into steadier trends later, suggesting speculation gave way to genuine utility and stabilized growth. All coins follow similar arcs - birth pangs before equilibrium. The analysis confirms Dash echoed the broader cryptocurrency market environment, seeing comparable stability phases and turbulence around key developments that likely impacted investor and user activity.

4. **LTC**: LTC's price trends show periods of price growth and decline. For instance, the transition from Change Point 13 (mean price approximately 51.17) to Change Point 14 (mean price approximately 87.71) shows a significant increase in the mean price. The highest kurtosis is observed at Change Point 20 (kurtosis approximately 2.20), indicating a period of increased market risk.

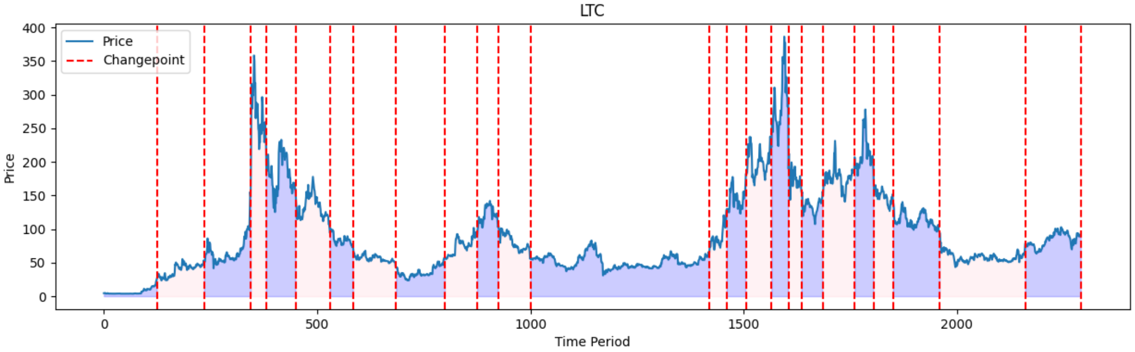


Figure 9 (d). Analytical depiction of Change Points in the Litecoin (LTC) time series.

It is observed in Table A9, in the Appendix A that throughout its development, the Litecoin market has exhibited a combination of stability and volatility, much like its predecessor cryptocurrencies. Periods such as 345 to 380, 525 to 585, and 1000 to 1420 demonstrate significant external changes associated with adoption and regulation, as evidenced by the exceptionally low U-statistics. Conversely, elevated U-statistics after initial unrest (e.g., 875 to 925, 1605 to 1635, 1850 to 1960) suggest a steadier approach to pricing that corresponds to growth that has become more stable. P-Values in Table A9 that are consistently significant across all change-points, further validate wider price fluctuations between the identified change points.

5. **XRP**: XRP's price trends also show non-linear behavior. For instance, the transition from Change Point 10 (mean price approximately 0.2363) to Change Point 11 (mean price approximately 0.5730) shows a significant increase in the mean price. The highest kurtosis is observed at Change Point 12 (kurtosis approximately 6.46), indicating a period of increased likelihood of extreme price changes.

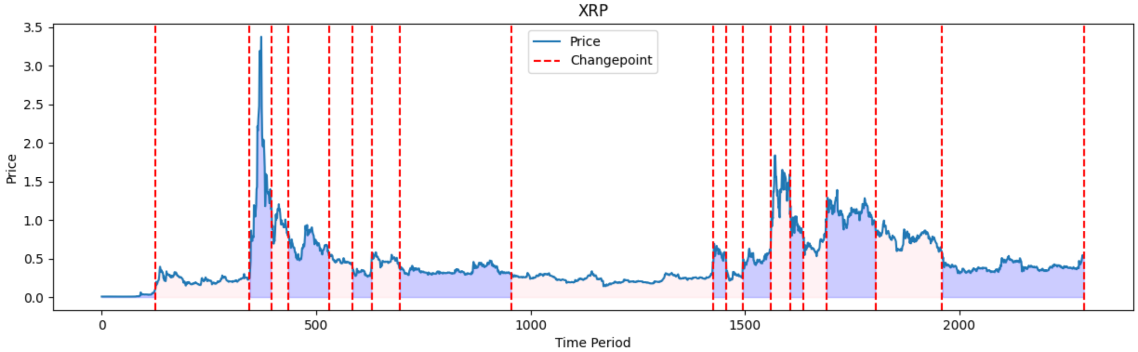


Figure 9 (e). Analytical depiction of Change Points in the Ethereum time series.

Table A10 illustrates the initial volatility of Ripple, which was subsequently followed by periods of stability interspersed with price fluctuations. This pattern mirrors well-established cryptocurrencies and substantiates Ripple's integration into the cryptocurrency market's evolution.

Ripple echoes volatility and steadiness patterns exhibited by other predecessor cryptocurrencies through its market progression. Extremely low U-Statistics like 0 between change points 1425-1455 and 1560-1605 indicate distributional divergence linked to external events, news, and shifting market adoption. Conversely, intervals such as 695-955, 1605-1635, and 1805-1960 display elevated U-Stats, pointing to pricing consistency and stabilized growth. The consistent P-value significance reinforces substantial intervals between the change points. This trend of volatility giving way to equilibrium mirrors analysis for Bitcoin, Ethereum, and Litecoin, suggesting similar early turbulence before broader integration. The findings confirm Ripple's position embedded in the wider cryptocurrency domain.

In summary, our analysis of the data presented in the above tables and figures reveals following key trends in the exchange prices of Bitcoin and Altcoins:

1. **Non-Linear Price Trends**: The price trends of these cryptocurrencies are non-linear, with distinct periods of price growth and decline. This is evidenced by the shifts in mean prices across the different segments. For example, Bitcoin's transition from Change point 1 to Change point 2 shows a significant increase in the mean price, indicating a bullish phase, while the transition from Change point 9 to Change point 10 shows a significant decrease, indicating a bearish phase. Similar non-linear trends can be observed in the other cryptocurrencies.
2. **Presence of Extreme Price Changes**: The data exhibits periods of leptokurtosis, indicating a higher probability of extreme price changes. This is evidenced by the high positive kurtosis values in certain segments, such as Bitcoin's Change point 14, suggesting increased market risk during these periods. Similar periods of leptokurtosis can be observed in the other cryptocurrencies, indicating the presence of fat tails in the price distribution.
3. **Asymmetry in Price Changes**: The data also exhibits periods of skewness, indicating an asymmetry in the distribution of price changes. For instance, the positive skewness in Bitcoin's Change point 2 suggests a higher probability of large price increases during this period, while the negative skewness in Change point 9 suggests a higher probability of large price decreases. Similar periods of skewness can be observed in the other cryptocurrencies, indicating that price changes are not symmetrically distributed.
4. **Significant Fluctuations in Volatility**: The standard deviation, a measure of price volatility, shows significant fluctuations across different segments. For example, Bitcoin's transition from Change point 8 to Change point 9 not only indicates a bearish phase with a decrease in the mean price, but also a decrease in volatility, suggesting a stabilization of prices after a period of high volatility. Conversely, the transition from Change point 6 to Change point 7 shows a bullish phase with an increase in both the mean price and volatility, indicating a period of rapid price growth. Similar fluctuations in volatility can be observed in the other cryptocurrencies, underscoring the inherent risk and uncertainty in the cryptocurrency market.
5. **Skewness and Kurtosis in Price Trends**: Bitcoin's Change point 15 has a high kurtosis of 2.9473, indicating a leptokurtic distribution and suggesting increased market risk due to more frequent extreme price changes. Skewness values, like the positive 0.7941 at Change point 2 and negative -0.3420 at Change point 9, show asymmetric price changes, with higher probabilities of large price increases and decreases, respectively. Similar patterns are seen in all other cryptocurrencies. The skewness and kurtosis values across all change points suggest a dynamic and volatile cryptocurrency market with periods of both bullish and bearish trends, and varying volatility levels.

In summary, the analysis reveals the non-linear, volatile, and asymmetric nature of the price trends of these cryptocurrencies, along with the significant fluctuations in volatility. These insights can be valuable for understanding the dynamics of the cryptocurrency market and potentially for informing investment decisions.

4.1.1.1.2 Results for Multifractal Detrended Fluctuation Analysis

The Hurst exponent is a statistical measure for determining the long-term memory of a financial time series. A Hurst exponent value of H = 0.5 indicates a random walk and the absence of long-term memory. In contrast, a value of H > 0.5 indicates positive long-term memory, which is frequently characterized by trend-following behavior. A value of H = 0.5, on the other hand, indicates negative long-term memory, typically associated with mean-reverting behavior. Figure 9(f) compares the Hurst exponents for Bitcoin and other Altcoins over daily and weekly timescales.

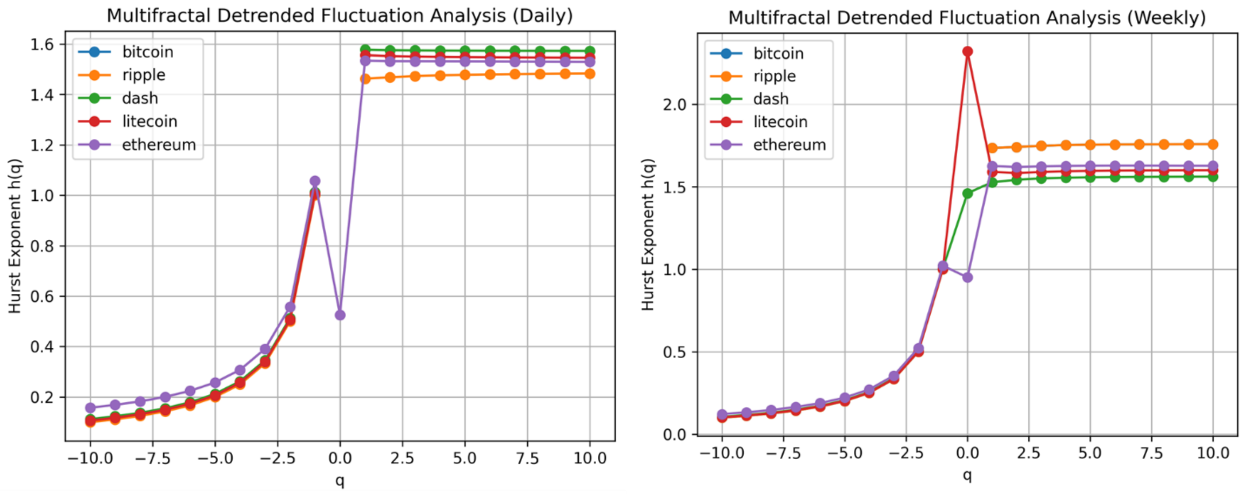


Figure 9 (f). Multifractal Detrended Fluctuation Analysis (MFDFA) of Bitcoin and Altcoins. The plots illustrate the generalized Hurst exponent, h(q), against the range of q values from -10 to 10 for both daily and weekly data. Each line represents a different cryptocurrency, illustrating the multifractal nature of their price characteristics over the Interval 4.

The generalized Hurst exponent, h(q), is used to probe the temporal correlations embedded within the price series. For all five-time series, several patterns can be observed:

i. **Consistent Rise for Negative q Values:** For $q<0$, the h(q) shows a consistent and sharp rise until the q values are very close to 0. This pattern is observed in both daily and weekly scales. This suggests that h(q) captures the large variations observed in the time series for all cryptocurrencies, which is indicative of persistent behavior (i.e., a positive correlation is exhibited between historical and future prices).

ii. **Sharp Rise:** It can also be observed that there is a steeper ascent in h(q) for q values transitioning from negative to positive than it was observed earlier. As q transitions from negative to positive, the ascent becomes more steeper which indicates that the MFDFA starts to emphasize larger fluctuations in the series. The increasing trend suggests that these fluctuations exhibit persistent behavior, indicative of long-range correlations or a long-memory process. Conversely, for q>1, the analysis gives more weight to larger fluctuations. The persistence here which is observed in all five-time series implies that both minor and significant price changes in the cryptocurrencies are temporally correlated.

iii. **Dip at q=0:** A noticeable dip in $h(q)$ at $q=0$ for all cryptocurrencies indicates a multifractal nature in the price series. Specifically, the dip suggests that the scaling behavior of the series, when all fluctuations are treated equally (i.e., at $q=0$), is distinct from the scaling behavior observed for other q values. This deviation is symbolic of a series that exhibits a range of scaling behaviors, a defining characteristic of multifractality.

iv. **Persistence for q>1 and Uniform Behavior Across Cryptocurrencies:** The relatively flat behavior of h(q) for q>1 across all cryptocurrencies suggests that the larger fluctuations in the price series, which are emphasized in this regime, exhibit similar scaling behavior. However, the values of hurst exponents vary significantly for daily and weekly scales. This implies the multifractal nature of cryptocurrency time series. The fluctuations may be attributed to a wide range of underlying factors. Some common factors that can influence the cryptocurrency market encompass a wide range of factors, including Blockchain information, macroeconomic factors, technical indicators, and other geopolitical events that influence cryptocurrency prices.

Overall, it would be safe to imply from the above analysis that Hurst exponents varied significantly on daily and weekly scales and there are distinct patterns in the temporal correlations of the price series for all five coins. The consistent rise of h(q) for negative q values in both daily and weekly scales suggests that the exponent captures significant variations in the time series for all examined cryptocurrencies. As the q values transition from negative to positive, a steeper ascent in h(q) is observed, emphasizing the larger fluctuations in the series, and indicating persistent behavior with long-range correlations. The noticeable dip at q=0 across all cryptocurrencies signifies the multifractal nature of the price series, with a distinct scaling behavior when all fluctuations are treated equally. Furthermore, the plateau observed for q>1 across all cryptocurrencies indicates that larger fluctuations in the price series have similar scaling behaviors, regardless of the daily or weekly scale. This behavior, combined with the variations in Hurst exponents between daily and weekly scales, confirms the multifractal nature of cryptocurrency time series. These observations emphasize the need for a sophisticated forecasting framework to capture and account for the multifaceted nature of cryptocurrency price movements. The above findings of persistence and multifractality are consistent with past studies such as (Caporale et al., 2018), (Eom et al., 2019) and (Wątorek et al., 2021).

4.1.1.2 Forecasting Results

In this section, we present the mean results from 20 runs of the proposed Deep learning models, including ANN, BiLSTM, and CNN-BiLSTM, each applied to different sets of indicators: Fundamental Indicators (FI), Technical Indicators (TI), and a combination of both Fundamental and Technical Indicators (FI+TI). We further conduct a comparison of our results with the benchmarks in the past literature to evaluate the performance of our models. In Appendix A, Tables A1 to A5 present the optimal hyperparameter values determined for each model by Bayesian Optimization for each coin. These values are specifically for Interval 4, which contains the latest data.

4.1.1.2.1 Comparative Analysis of Predictive Models

We conducted an ablation analysis to evaluate the performance of various models, using different combinations of input datasets consisting of fundamental indicators, technical indicators, and combined fundamental and technical indicators in predicting the closing prices of Bitcoin, Ethereum, and other Altcoins across different prediction horizons, from the next day to 7 days.

In Tables 16 through 20, we present the findings for the next day's closing price predictions for Bitcoin, Ethereum, Ripple, Litecoin, and DASH. Subsequently, Tables 21 to 25 provide performance metrics for forecasting periods spanning 3 days, 5 days, and 7 days. For a more comprehensive understanding, Figures 10 to 14 illustrate a comparison between the actual and predicted closing prices for each coin during Interval 4. Additionally, Figures 15 through 19 display the loss metrics for the 3-day, 5-day, and 7-day forecast periods for each respective coin.

Table 16. Mean results of 20 runs for next day closing price prediction of **Bitcoin** (BTC) for all four intervals

		CNN-BiLSTM			**BiLSTM**			**ANN**		
	Metric	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**
Interval 1	**MAE**	4.2705	**0.3771**	2.3075	5.5516	2.5415	4.4854	6.1103	1.9247	5.6095
	MAPE	0.7807	**0.0655**	0.3920	1.0144	0.4291	0.8192	1.1182	0.3257	1.0271
	RMSE	4.3943	**0.5201**	3.4255	5.7135	3.7635	4.6154	6.2855	2.8535	5.7721
Interval 2	**MAE**	2.3205	**0.8971**	8.1387	3.0169	2.4375	8.9505	3.3215	3.0485	6.7795
	MAPE	0.2085	**0.0918**	0.7865	0.2733	0.2216	0.8645	0.2992	0.2793	0.6565
	RMSE	3.2631	**1.0465**	9.0352	4.2445	3.4255	9.9385	4.6677	4.2835	7.5278
Interval 3	**MAE**	78.9609	**18.9521**	134.5681	111.2026	89.8433	160.3615	122.3235	112.3073	121.4853
	MAPE	1.0986	**0.1861**	1.7646	1.5477	1.2486	2.0995	1.7033	1.5644	1.5925
	RMSE	119.2203	**52.5120**	233.7843	167.9015	135.5835	278.5911	184.6919	169.4818	211.0553
Interval 4	**MAE**	340.1411	**145.8120**	227.0201	529.0417	413.4718	591.5748	1279.9001	633.3301	776.5103
	MAPE	3.9812	**2.0804**	4.1146	2.7510	4.5117	3.3959	5.0824	2.8110	4.9255
	RMSE	324.5110	**183.5900**	225.7221	423.0568	346.3945	588.1511	1088.3140	537.1642	632.4784

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error. FI denotes Fundamental Indicators, TI represents Technical Indicators, and FI+TI refers to the input dataset containing both fundamental and technical indicators.

Table 17. Mean results of 20 runs for next day closing price prediction of **Ethereum** (ETH) for all four intervals

		CNN-BiLSTM			BiLSTM			ANN		
	Metric	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**
Interval 1	**MAE**	2.7103	**2.4584**	2.7576	3.0726	2.7759	2.9349	4.1239	3.2305	3.1562
	MAPE	0.0735	**0.0691**	0.0817	0.1033	0.0981	0.1191	0.1595	0.1134	0.1716
	RMSE	1.3935	**1.0952**	1.4742	1.6809	1.5062	1.6564	2.2552	1.7185	1.7929
Interval 2	**MAE**	5.8921	**5.3443**	5.9948	6.6795	6.0346	6.3803	8.9649	7.0229	6.8613
	MAPE	0.1414	0.1328	0.1571	0.1987	**0.1316**	0.2290	0.3067	0.2180	0.3299
	RMSE	6.9674	**5.4760**	7.3710	8.4044	7.5309	8.2821	11.2758	8.5924	8.9647
Interval 3	**MAE**	58.9206	52.4429	59.9476	66.7952	**52.3461**	63.8028	89.6492	70.2293	68.6127
	MAPE	1.4138	1.3285	1.5707	1.9872	1.8862	2.2904	3.0668	2.1804	**1.2995**
	RMSE	69.6745	**54.7602**	73.7095	84.0443	75.3090	82.8211	112.7583	85.9241	89.6471
Interval 4	**MAE**	98.2010	**89.0715**	99.9126	111.3253	100.5768	106.3380	149.4154	117.0489	114.3545
	MAPE	2.3564	**2.2141**	2.6179	3.3120	3.1436	3.8173	5.1113	3.6340	5.4991
	RMSE	116.1241	**91.2670**	122.8492	140.0739	125.5150	138.0352	187.9305	143.2069	149.4119

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error.

Table 18. Mean results of 20 runs for next day closing price prediction of **Ripple** (XRP) for all four intervals

		CNN-BiLSTM			BiLSTM			ANN		
	Metric	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**
Interval 1	**MAE**	0.0023	**0.0014**	0.0023	0.0044	0.0082	0.0096	0.0205	0.0439	0.0285
	MAPE	0.3684	**0.3307**	0.6649	1.2695	1.0959	1.5145	0.8879	0.8740	1.7848
	RMSE	0.0019	**0.0012**	0.0019	0.0035	0.0066	0.0078	0.0166	0.0356	0.0231
Interval 2	**MAE**	0.0036	**0.0022**	0.0036	0.0068	0.0127	0.0150	0.0319	0.0682	0.0443
	MAPE	0.6581	**0.5907**	1.1876	2.2676	1.9574	2.7052	1.5859	1.5612	3.1880
	RMSE	0.0030	**0.0018**	0.0029	0.0055	0.0103	0.0121	0.0259	0.0553	0.0359
Interval 3	**MAE**	0.0057	**0.0035**	0.0055	0.0106	0.0197	0.0233	0.0496	0.1060	0.0688
	MAPE	0.5439	**0.4882**	0.9815	1.8740	1.6177	2.2357	1.3107	1.2903	2.6348
	RMSE	0.0046	0.0028	0.0045	0.0086	**0.0025**	0.0189	0.0402	0.0859	0.0557
Interval 4	**MAE**	0.0088	**0.0054**	0.0086	0.0164	0.0306	0.0362	0.0771	0.1647	0.1069
	MAPE	0.4386	**0.3937**	0.7915	1.5113	1.3046	1.8030	1.0570	1.0405	2.1248
	RMSE	0.0083	**0.0059**	0.0071	0.0088	0.0166	0.0378	0.0877	0.1473	0.1975

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error.

Table 19. Mean results of 20 runs for next day closing price prediction of **Litecoin** (LTC) for all four intervals

		CNN-BiLSTM			BiLSTM			ANN		
	Metric	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**
Interval 1	**MAE**	0.9052	**0.8096**	0.9565	1.0069	0.9515	1.0569	1.1022	1.0842	1.1263
	MAPE	2.0035	**1.8060**	2.1026	2.3013	2.2089	2.4042	2.6089	2.8055	3.0021
	RMSE	0.4041	**0.3571**	0.4523	0.5038	0.4864	0.5288	0.6017	0.5857	0.6294
Interval 2	**MAE**	0.9295	**0.8235**	0.9764	1.0238	0.9724	1.0774	1.1250	1.1005	1.1438
	MAPE	2.1061	**1.9026**	2.2172	2.4065	2.3010	2.5004	2.7152	2.9043	3.1059
	RMSE	0.4290	**0.3731**	0.4790	0.5286	0.5005	0.5403	0.6244	0.6007	0.6437
Interval 3	**MAE**	0.9457	0.8469	0.9957	0.9435	**0.8185**	1.0946	1.1434	1.1267	1.1632
	MAPE	2.2057	**2.0025**	2.3072	2.5069	2.4078	2.6017	2.8252	3.2028	3.2773
	RMSE	0.4488	**0.3989**	0.4907	0.5404	0.5296	0.5672	0.6448	0.6263	0.6692
Interval 4	**MAE**	0.9664	**0.8602**	1.0162	1.0615	1.0104	1.1165	1.1646	1.1471	1.1878
	MAPE	1.3932	**1.2054**	2.4173	2.6565	2.5295	2.7023	2.9009	3.1032	3.3022
	RMSE	0.1611	**0.1209**	0.1545	0.5653	0.5419	0.5875	0.6641	0.6417	0.6887

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error.

Table 20. Mean results of 20 runs for next day closing price prediction of **DASH** coin for all four intervals

		CNN-BiLSTM			BiLSTM			ANN		
	Metric	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**	**FI**	**FI+TI**	**TI**
Interval 1	**MAE**	0.0584	0.0497	0.0643	0.0744	**0.0351**	0.0839	2.0006	1.9827	2.0238
	MAPE	1.0035	**0.9083**	1.1040	1.2010	1.1563	1.2502	2.5063	2.4894	2.5228
	RMSE	0.1064	**0.0949**	0.1177	0.1239	0.1201	0.1351	2.7066	2.6896	2.7281
Interval 2	**MAE**	0.0527	**0.0426**	0.0691	0.0748	0.0757	0.0843	2.0291	2.0053	2.0455
	MAPE	1.0519	0.9519	1.1560	1.2065	**0.9486**	1.3044	2.5586	2.5359	2.5752
	RMSE	0.1085	**0.0938**	0.1185	0.1276	0.1206	0.1367	2.7279	2.7021	2.7455
Interval 3	**MAE**	0.0567	**0.0488**	0.0650	0.0796	0.0746	0.0831	2.0453	2.0246	2.0647
	MAPE	1.1034	1.0061	1.2033	**1.0024**	1.2502	1.3540	2.6013	2.5884	2.6246
	RMSE	0.1057	**0.0981**	0.1186	0.1233	0.1254	0.1307	2.7407	2.7218	2.7697

Interval 4	**MAE**	0.0681	**0.0551**	0.0761	0.0878	0.0770	0.0895	2.0668	2.0496	2.0824
	MAPE	1.1556	**1.0581**	1.2555	1.3545	1.3066	1.4097	2.6559	2.6305	2.6707
	RMSE	0.1149	**0.1079**	0.1330	0.1369	0.1245	0.1347	2.7633	2.7482	2.7881

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error.

Table 16 provides a comprehensive overview of the mean results from 20 runs for predicting the next day's closing price of Bitcoin (BTC) across all four intervals. The results are segmented by three models: CNN-BiLSTM, BiLSTM, and ANN, and further divided based on the type of input dataset used, namely Fundamental Indicators (FI), Technical Indicators (TI), and a combination of both (FI+TI).

For Interval 1, the CNN-BiLSTM model with FI+TI as input exhibits the most accurate predictions, as evidenced by its lowest MAE, MAPE, and RMSE values of 0.3771, 0.0655%, and 0.5201, respectively. This trend of superior performance by the CNN-BiLSTM model with FI+TI continues in Interval 2, where it achieves MAE, MAPE, and RMSE values of 0.8971, 0.0918%, and 1.0465, respectively. During Interval 3, the CNN-BiLSTM model with FI+TI demonstrates superior performance by achieving the lowest values for MAE, MAPE, and RMSE, which are 18.9521, 0.1861%, and 52.5120, respectively. For Interval 4, while the CNN-BiLSTM model with FI+TI achieves the lowest MAE, RMSE, and MAPE values of 145.8120 and 183.5900, and 2.0804% respectively.

Interval 1 consistently exhibits the best performance across all models. This superior accuracy can be attributed to the nascent market stage during the Pre-Mount Gox Crash Period. The lack of external shocks and the premature nature of the market likely contributed to more predictable price movements. In contrast, Interval 4 encompasses both the boom-and-bust phases of the market and represents a more complex and representative period in the current context. The increased data from this interval captures the market's reactions to various macro-economic and speculatory events. These events introduced greater volatility and unpredictability, making accurate forecasting more challenging. However, the CNN-BiLSTM model, especially when FI and TI are used together, shows resilience and adaptability in capturing these complex dynamics, as shown by its performance consistently being good across intervals. This model's ability to analyze both financial and textual indicators (FI) and textual indicators (TI) allows it to capture a comprehensive market view.

Table 17 presents the performance metrics for the next day's closing price of Ethereum (ETH) across all four intervals. For Interval 1, the CNN-BiLSTM model with FI+TI as input demonstrates the most accurate predictions, as evidenced by its lowest MAE, MAPE, and RMSE values of 2.4584, 0.0691%, and 1.0952, respectively. This trend of superior performance by the CNN-BiLSTM model with FI+TI continues in Interval 2, where it achieves MAE, and RMSE values of 5.3443, and 5.4760, respectively. However, the BiLSTM model with FI+TI performs marginally better in MAPE with a value of 0.1328%. In Interval 3, the CNN-BiLSTM model with FI+TI achieves the lowest RMSE of 54.7602. However, ANN Model with Technical Indicators registers the lowest MAPE value of 1.2995% and BiLSTM with FI+TI performs best in terms of MAE with a value of 52.3467. For Interval 4, the CNN-BiLSTM model with FI+TI achieves the lowest MAE, RMSE, and MAPE values of 89.0715, 91.2670, and 2.2141% respectively.

From the results, it can be observed that Ethereum exhibits similar trends in performance metrics as seen in Table 16 for Bitcoin. The results observed in Table 16 for Bitcoin demonstrate a similar pattern to the progression from Interval 1 to 4, indicating the consistent behavior of these models across both Bitcoin and Ethereum.

Table 18 provides Ripple (XRP) performance metrics across all four intervals. It is evident from the results that the CNN-BiLSTM model with FI+TI as the input dataset is the best-performing model for Ethereum (ETH) predictions. This suggests that integrating fundamental and technical indicators offers a more robust predictive capability. A similar pattern can be seen in Tables 19 and 20, where the CNN-BiLSTM model with FI+TI input remains the top performer for both Litecoin and DASH coin predictions. When the performance metrics of Tables 16-20 are compared, it can be observed that Bitcoin's models in Table 16 show higher values of RMSE, MAPE, and MAE compared to the Altcoins. Bitcoin's exposure to a broader array of internal and external factors complicates its forecasting relative to other cryptocurrencies.

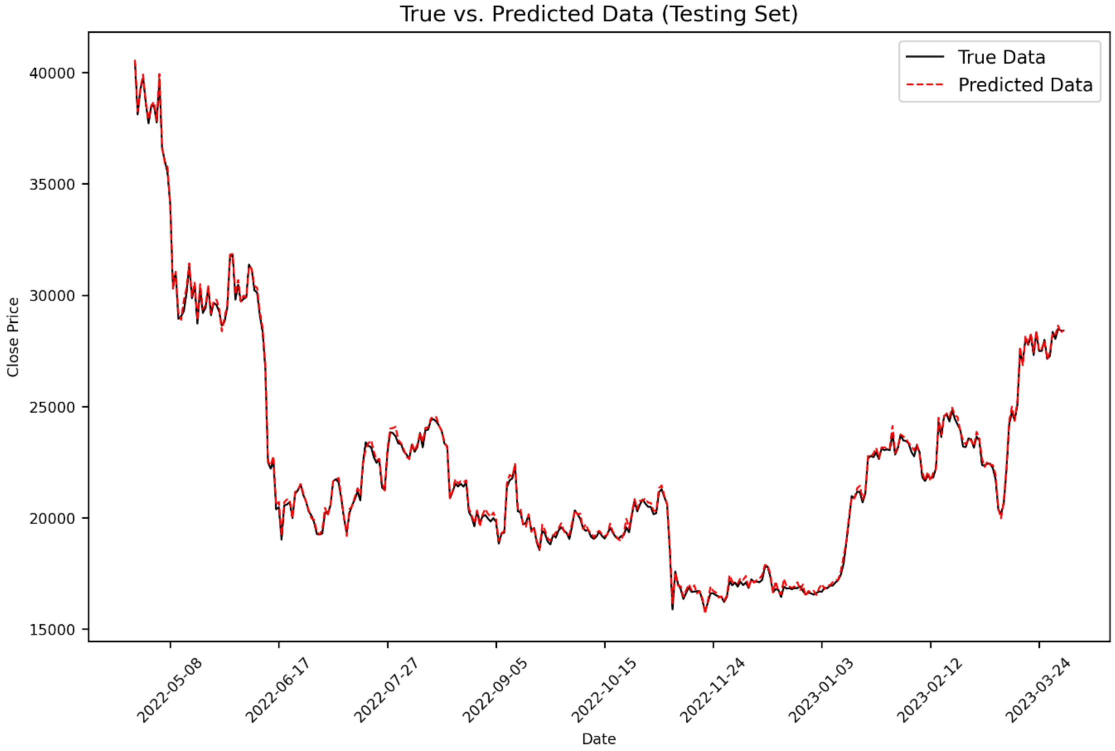


Figure 10. Comparison of predicted and actual closing prices of Bitcoin for Interval 4 using the best-performing model (testing dataset).

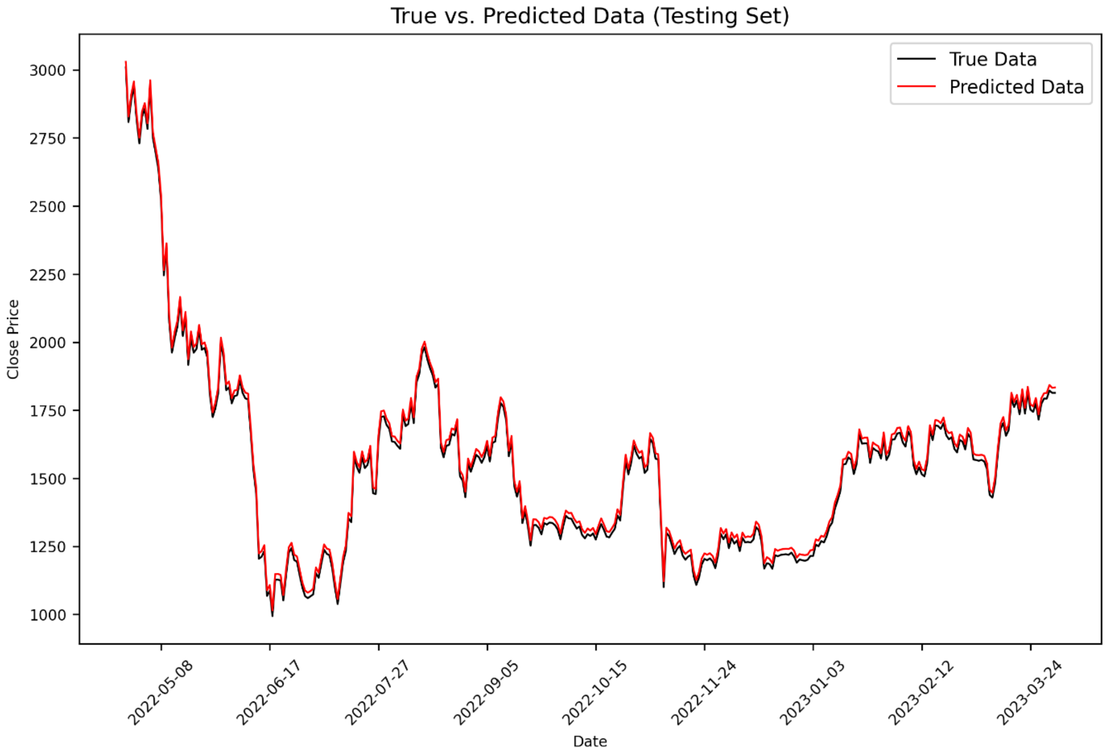


Figure 11. Comparison of predicted and actual closing prices of Ethereum for Interval 4 using the best-performing model (testing dataset).

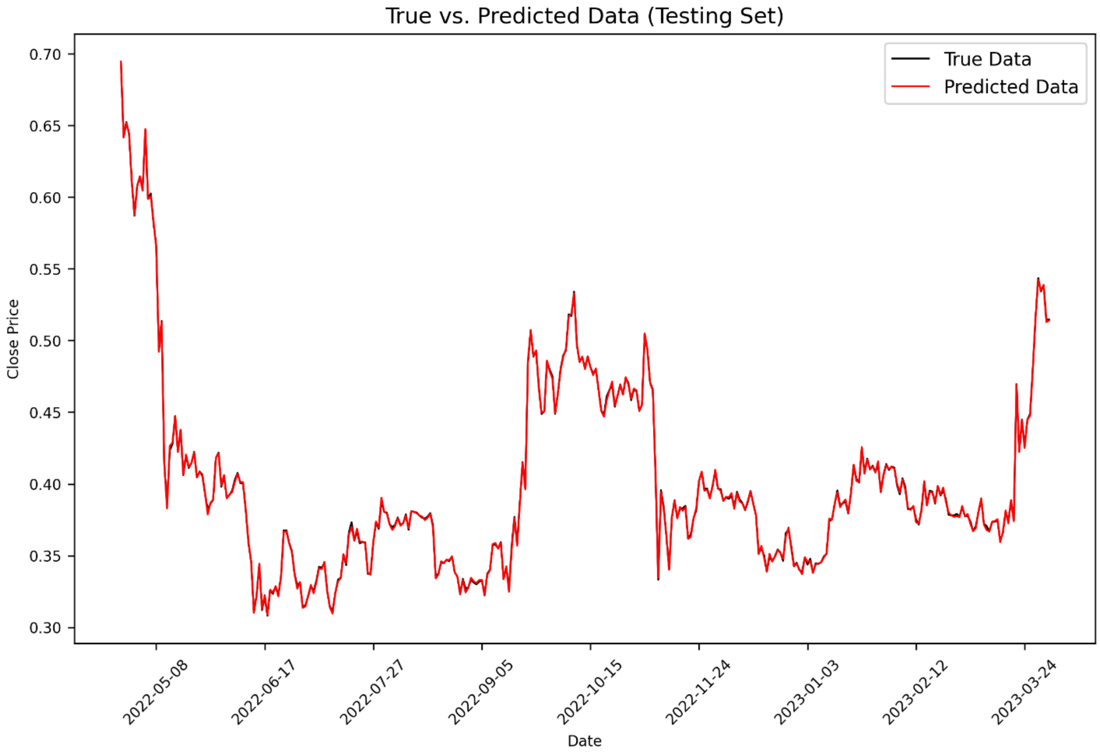


Figure 12. Comparison of predicted and actual closing prices of Ripple for Interval 4 using the best-performing model (testing dataset).

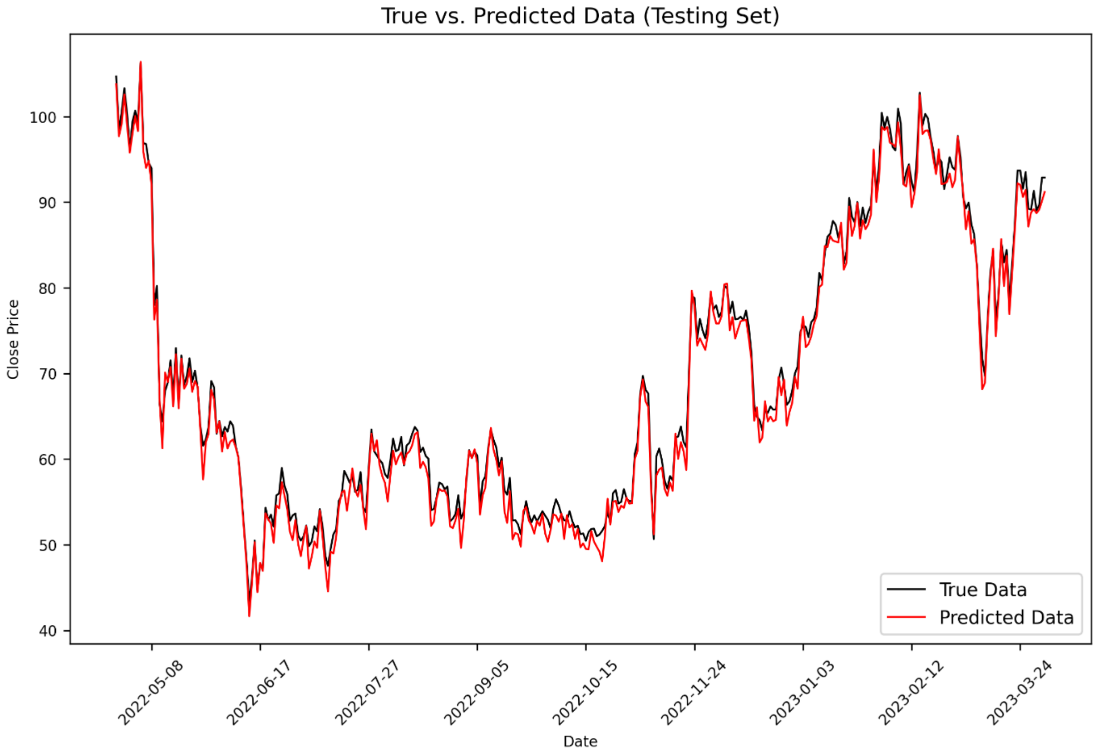


Figure 13. Comparison of predicted and actual closing prices of Litecoin for Interval 4 using the best-performing model (testing dataset).

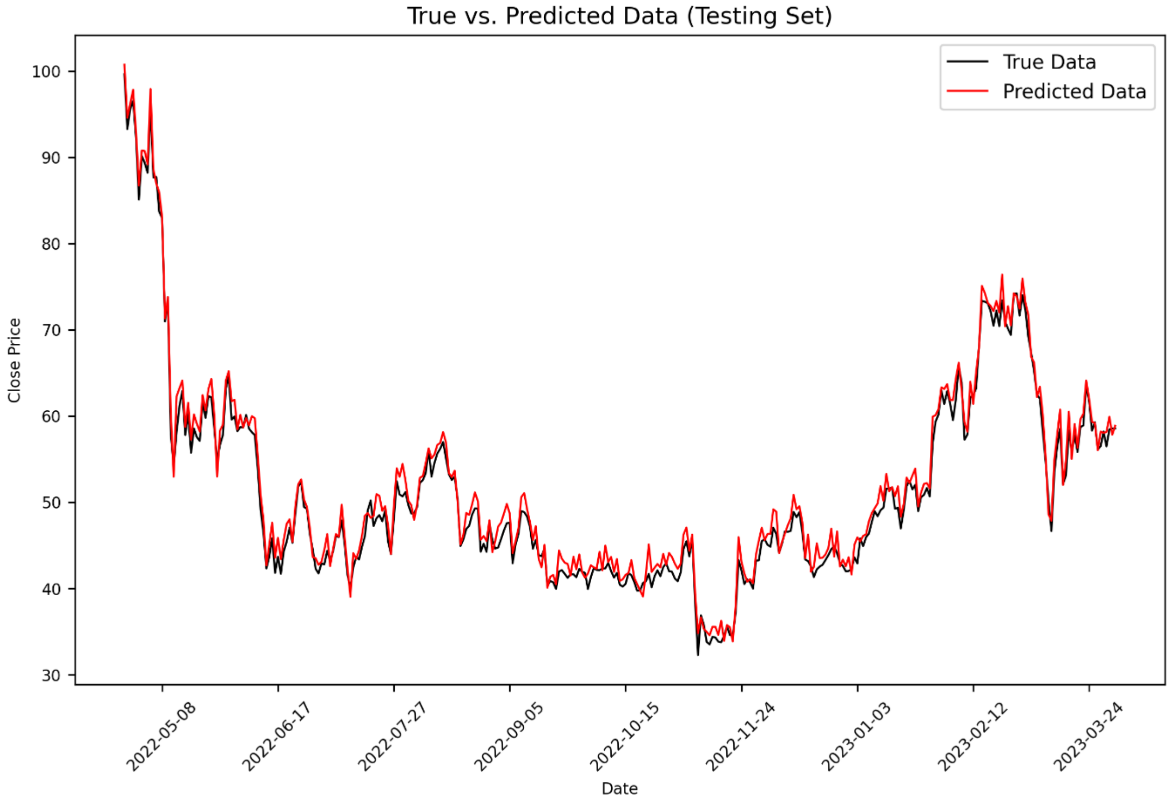


Figure 14. Comparison of predicted and actual closing prices of DASH for Interval 4 using the best-performing model (testing dataset).

Figures 10-14 illustrate a comparison between the predicted and actual closing prices for Bitcoin and the examined altcoins over Interval 4 using the best-performing modeled. Visually, a high degree of alignment is evident between the forecasted and true market prices across the assets, even during periods of heightened volatility. The sustained precision across coins and time affirms the model's capability to reliably forecast prices amid cryptocurrency complexity.

Table 21. The mean forecast results for **Bitcoin's** closing price over 'N' days across 20 runs. This table displays the results of Interval 4, the most recent interval.

	3 days			5 days			7 days		
Metrics	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**	**MAE**	**RMSE**	**SMAPE**	**MAE**
ANN (FI)	1390.12	4.22	1158.14	1515.65	4.51	1244.41	1598.38	5.18	1299.43
ANN (TI)	834.21	4.27	1151.53	1514.616	4.55	1244.42	1601.34	5.13	1303.12
ANN (FI+TI)	688.71	2.01	677.69	846.24	4.45	694.45	903.67	5.34	736.16
BiLSTM (FI)	628.87	2.64	567.67	716.74	4.31	601.28	819.73	4.75	660.17
BiLSTM (TI)	457.32	3.17	488.02	647.47	3.08	502.33	711.30	4.25	568.92
BiLSTM (FI+TI)	846.59	1.68	370.29	498.52	2.55	410.04	512.86	4.06	415.57
CNN-BiLSTM (FI)	420.68	2.02	843.71	872.08	3.04	863.01	966.74	3.82	963.63
CNN-BiLSTM (TI)	275.65	2.09	406.01	401.81	2.61	384.54	525.51	3.89	515.91
CNN-BiLSTM (FI+TI)	**142.24**	**0.88**	**268.01**	**292.23**	**1.67**	**299.72**	**380.91**	**3.41**	**381.6**

Source: Self Compilation.

The values highlighted in bold represent the lowest forecasting error. FI denotes Fundamental Indicators, TI represents Technical Indicators, and FI+TI refers to the input dataset containing both fundamental and technical indicators.

Table 22. The mean forecast results for **Ethereum's** closing price over 'N' days across 20 runs. This table displays the results of Interval 4, the most recent interval.

	3 days			5 days			7 days		
Metrics	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**	**MAE**
ANN (FI)	1382.75	2.10	1151.13	1514.16	3.43	1247.84	1595.35	4.30	1301.62
ANN (TI)	832.34	2.09	682.86	843.35	3.19	688.33	906.96	4.41	736.97
ANN (FI+TI)	681.04	2.88	575.52	716.99	3.05	598.71	823.68	4.72	661.16
BiLSTM (FI)	625.26	2.65	489.82	638.76	2.19	504.26	713.91	4.98	562.66
BiLSTM (TI)	461.20	3.29	379.85	498.50	4.16	404.90	515.49	4.61	411.21
BiLSTM (FI+TI)	851.72	2.09	848.31	866.42	3.50	863.73	975.78	4.44	970.75
CNN-BiLSTM (FI)	414.33	1.53	407.39	403.07	1.88	383.39	523.96	2.82	509.61
CNN-BiLSTM (TI)	273.75	1.51	273.31	300.94	**1.73**	299.96	379.08	**1.34**	380.78
CNN-BiLSTM (FI+TI)	**136.5**	**0.87**	**144.79**	**154.26**	1.76	**150.36**	**249.55**	1.77	**247.41**

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error.

FI denotes Fundamental Indicators, TI represents Technical Indicators, and FI+TI refers to the input dataset containing both fundamental and technical indicators.

Table 23. The mean forecast results for **Ripples'** closing price over 'N' days across 20 runs. This table displays the results of Interval 4, the most recent interval.

	3 days			5 days			7 days		
Metrics	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**
ANN (FI)	2.6248	3.3730	3.5094	3.2285	3.8115	4.3166	4.0680	4.4213	5.4389
ANN (FI+TI)	2.4210	3.1355	3.1803	3.0008	3.2914	3.7521	3.0308	3.1014	3.9121
ANN (TI)	2.6655	3.4185	3.5409	3.2785	3.8629	4.3553	3.5385	4.1229	4.6153
BiLSTM (FI)	0.1115	1.6145	0.1725	0.1372	1.8244	0.2122	0.3972	2.0844	0.5722
BiLSTM (FI+TI)	0.0832	1.5810	0.1506	0.1146	**1.1549**	0.1793	0.3746	1.9149	0.4393
BiLSTM (TI)	0.1128	1.6697	0.1711	0.1387	1.8868	0.2104	0.3987	2.1468	0.5704
CNN-BiLSTM (FI)	0.0851	1.4156	0.1432	0.1047	1.5996	0.1762	0.3647	1.8596	0.5562
CNN-BiLSTM (FI+TI)	**0.0378**	**1.1861**	**0.1087**	**0.0834**	1.2276	**0.1112**	**0.2434**	**1.1476**	0.3012
CNN-BiLSTM (TI)	1.8264	1.5694	0.1649	2.2465	1.7734	0.2029	2.8306	2.2345	**0.2779**

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error.

FI denotes Fundamental Indicators, TI represents Technical Indicators, and FI+TI refers to the input dataset containing both fundamental and technical indicators.

Table 24. The mean forecast results for **Litecoin**'s closing price over 'N' days across 20 runs. This table displays the results of Interval 4, the most recent interval

	3 days			**5 days**			**7 days**		
Metrics	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**
ANN (FI)	1.4674	3.6551	0.8368	1.8489	4.0938	1.0543	2.4036	4.6669	1.3706
ANN (FI+TI)	1.3880	3.7549	0.7765	1.7489	4.2055	0.9783	2.2735	4.6942	1.2718
ANN (TI)	1.4848	4.1278	0.8609	1.8708	4.6231	1.0847	2.4320	5.1703	1.4101
BiLSTM (FI)	1.3375	3.3472	0.7123	1.6852	3.7489	0.8975	2.1908	4.2737	1.1667
BiLSTM (FI+TI)	1.2226	3.0607	0.6557	1.5405	3.4280	0.8262	2.0026	3.9079	1.0740
BiLSTM (TI)	1.3956	3.3779	0.7344	1.7585	3.7832	0.9253	2.2860	4.3029	1.2029
CNN-BiLSTM (FI)	1.2177	1.7554	0.5810	1.5343	1.9661	0.7320	**1.4945**	**1.2513**	0.9517
CNN-BiLSTM (FI+TI)	**0.9408**	**1.4585**	**0.1603**	**1.1855**	**1.6336**	**0.1317**	1.5411	1.8923	**0.8342**
CNN-BiLSTM (TI)	1.2703	3.0216	0.6306	1.6005	3.3842	0.7946	2.0807	3.8280	1.0330

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error.

FI denotes Fundamental Indicators, TI represents Technical Indicators, and FI+TI refers to the input dataset containing both fundamental and technical indicators.

Table 25. The mean forecast results for **DASH** coin's closing price over 'N' days across 20 runs. This table displays the results of Interval 4, the most recent interval.

	3 days			**5 days**			**7 days**		
Metrics	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**
ANN (FI)	2.5628	3.2933	3.4265	3.4855	3.5568	4.5915	4.3220	4.0192	5.6935
ANN (FI+TI)	2.5415	3.2618	3.4078	3.4564	3.5228	4.5664	4.2860	3.9807	5.6623
ANN (TI)	2.5822	3.3117	3.4572	3.5118	3.5766	4.6327	4.3546	4.0416	5.7446
BiLSTM (FI)	0.1089	1.6796	0.1698	0.1481	1.8139	0.2275	0.1836	2.0498	0.2821
BiLSTM (FI+TI)	0.0955	1.6202	0.1544	0.1299	1.7498	0.2069	0.1610	1.9773	0.2565
BiLSTM (TI)	0.1110	1.7480	0.1670	0.1509	1.8879	0.2238	0.1872	2.1333	0.2775
CNN-BiLSTM (FI)	0.0844	1.4329	0.1425	0.1148	1.5476	0.1909	0.1424	1.7488	0.2367
CNN-BiLSTM (FI+TI)	**0.0683**	**1.3120**	**0.1338**	**0.0929**	**1.4170**	**0.1793**	**0.1152**	**1.6012**	**0.1223**
CNN-BiLSTM (TI)	0.0944	1.5568	0.1649	0.1283	1.6814	0.2210	0.1591	1.8999	0.2740

Source: Self Compilation. The values highlighted in bold represent the lowest forecasting error.

FI denotes Fundamental Indicators, TI represents Technical Indicators, and FI+TI refers to the input dataset containing both fundamental and technical indicators.

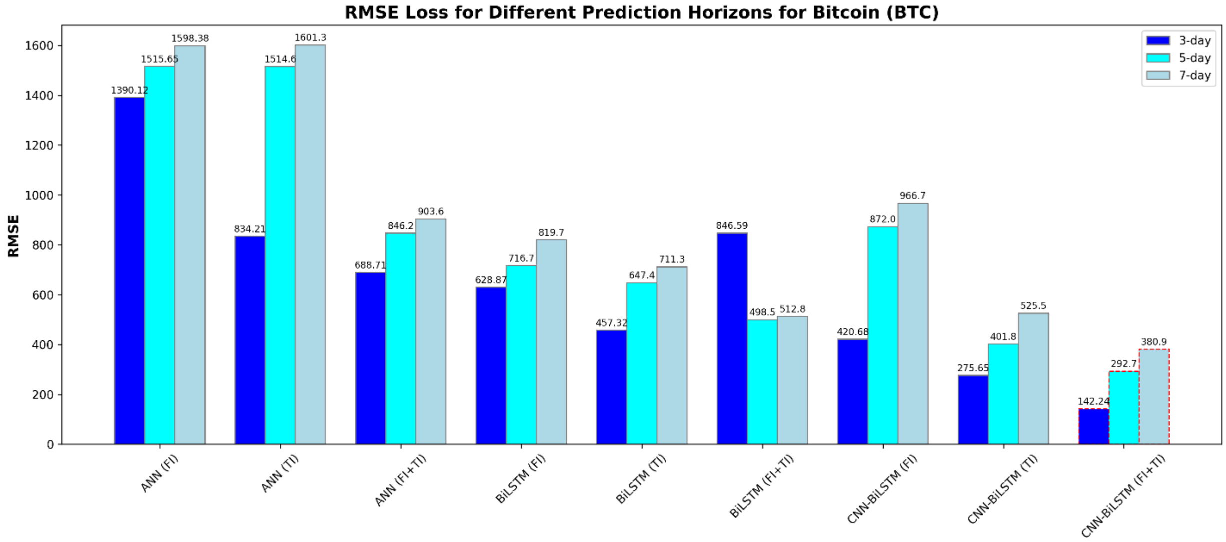


Figure 15. Model loss progression for Bitcoin over 3-day, 5-day, and 7-day forecast horizons (testing dataset).

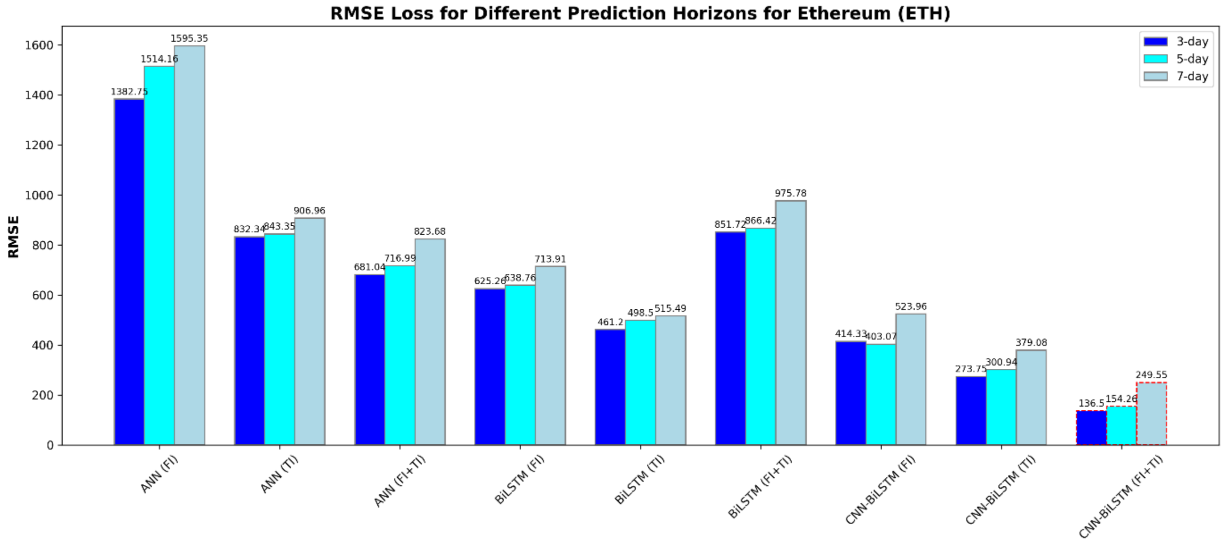


Figure 16. Model loss progression for Ethereum over 3-day, 5-day, and 7-day forecast horizons (testing dataset).

Tables 21 to 25 present the performance metrics for forecasting periods spanning 3 days, 5 days, and 7 days. It can be observed in Figure 15 and Table 21 that over a 3-day forecast, the CNN-BiLSTM model with both Fundamental and Technical Indicators (FI+TI) records the lowest RMSE value at 142.24. This trend is consistent over the 5-day and 7-day forecasts, with the model registering RMSE values of 292.23 and 380.91, respectively. This consistent performance across varying forecasting horizons indicates that the CNN-BiLSTM (FI+TI) model offers the most reliable predictions for Bitcoin's closing prices. A similar trend is observed in Ethereum's forecasting horizons in Table 22. The CNN-BiLSTM (FI+TI) model outperforms its counterparts across all horizons. Specifically, the model achieves RMSE values of 136.5, 154.26, and 249.55 for the 3-day, 5-day, and 7-day forecasts, respectively. This consistency in performance suggests that the CNN-BiLSTM (FI+TI) model is also the most effective for predicting Ethereum's closing prices.

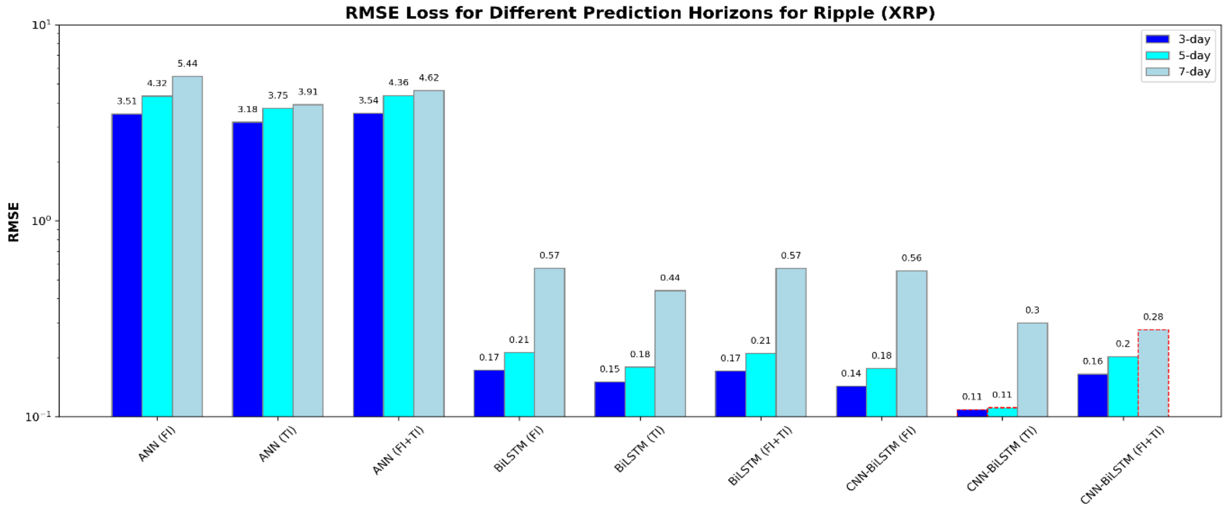


Figure 17. Model loss progression for Ripple over 3-day, 5-day, and 7-day forecast horizons (testing dataset).

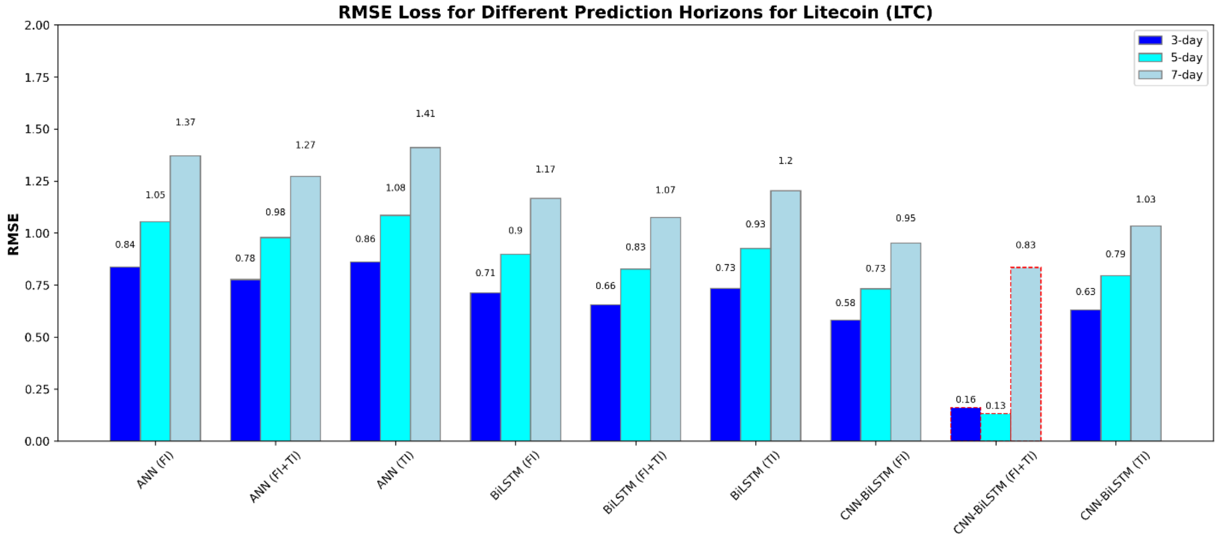


Figure 18. Model loss progression for Litecoin over 3-day, 5-day, and 7-day forecast horizons (testing dataset).

Tables 21-25 generally reveal that the CNN-BiLSTM model consistently records the lowest forecasting error values, especially when equipped with both Fundamental and Technical Indicators. This performance suggests that this model is particularly adept at predicting cryptocurrency closing prices over varying days. However, it is worth noting that while the CNN-BiLSTM model with FI+TI input consistently excels, there are scenarios where other models, such as the BiLSTM with either FI or TI, also perform commendably. This variability indicates that while the CNN-BiLSTM model is generally superior, other models might be suitable under specific conditions. Despite the differences in metrics across models and cryptocurrencies, the CNN-BiLSTM model with FI+TI input consistently outperforms other model-input combinations.

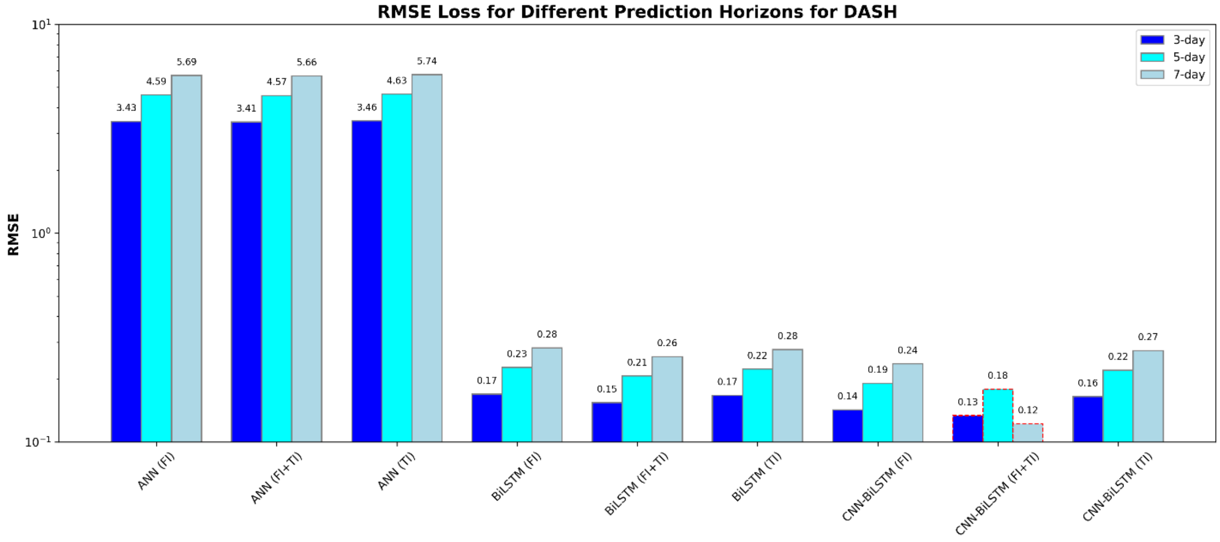


Figure 19. Model loss progression for DASH over 3-day, 5-day, and 7-day forecast horizons (testing dataset).

It can also be observed from the Figures and Tables 21-25 that ANNs are generally the worst-performing models. For Bitcoin, Ethereum, Ripple, Litecoin, and DASH coin, as detailed in Tables 21 through 25, ANN models typically register higher error metrics than their BiLSTM and CNN-BiLSTM counterparts. This tendency is especially evident when the ANN models use only Fundamental or Technical Indicators. In contrast, the more advanced BiLSTM and CNN-BiLSTM models, particularly those using both FI and TI, consistently offer better forecasting accuracy.

4.1.1.2.2 Comparing Our Results with Previous Literature

In this research, we critically assess the forecasting accuracy of our proposed models by comparing them with established benchmarks from previous literature in next-day and multi-step cryptocurrency price predictions. Our results indicate that our models consistently outperform the previous state-of-the-art works in most performance metrics across both next step and multi-step forecasting horizons.

1. **Comparison with Past Research for Bitcoin**

   Table 16 showcases the predictive capabilities of our CNN-BiLSTM model with composite Fundamental indicators + Technical indicators (FI+TI) for Bitcoin's next-day closing price across four intervals. For Interval 4, our CNN-BiLSTM model achieved an RMSE of 183.5900, an MAE of 145.8120, and a MAPE of 2.0804%. When juxtaposed against benchmarks from the literature, our model demonstrates superior performance in several instances. Specifically, it outperforms the RMSE values reported by (Aggarwal et al., 2020), which was calculated to be 386.70557, (Luo et al., 2022) with an RMSE of 244.303, (Altan et al., 2019) with an RMSE of 623.41, (Ye et al., 2022), with an RMSE of 525.533, (Ferdiansyah et al., 2023) with a staggering RMSE of 2343.2200, (El-Berawi et al., 2021) with an RMSE of 850.8, (Fleischer et al., 2022) with an RMSE of 1334.755, (Maciel et al., 2022b) with an RMSE of 1802.6325, and (P. Kumar et al., 2022) with an RMSE of 198.0065. In terms of MAE, our model surpassed (Luo et al., 2022), which reported an MAE of 167.331, Altan et al. (2019) with an MAE of 500.78, (Ye et al., 2022) with an MAE of 359.08, and (Maciel et al., 2022b) with an MAE of 1313.8272. Furthermore, our model's MAPE of 2.0804% was more accurate than (Altan et al., 2019) with a MAPE of 3.55%, (Ye et al., 2022) with a MAPE of 3.18%, (Ferdiansyah et al., 2023) with a MAPE of 4.0%, (El-Berawi et al., 2021) with a MAPE of 2.28%, and (Maciel et al., 2022b) with a MAPE of 2.8320%. This comparative analysis empirically demonstrates the enhanced predictive performance of our CNN-BiLSTM model (FI+TI) against several established benchmarks in the previous literature.

2. Comparison with Past Research for Ethereum

Table 17 presents the results of our predictive models for the next-day closing price prediction of Ethereum (ETH) across four intervals. The CNN-BiLSTM model with composite Technical + Fundamental indicators consistently demonstrates superior RMSE, MAE, and MAPE metrics performance. For instance, (El-Berawi et al., 2021) reported an RMSE of 82.03 for a similar period as our Interval 3, while our CNN-BiLSTM model achieved an RMSE of 54.7602 in the same interval. (Maciel et al., 2022b) documented an MAE of 112.5590, but our model reached an MAE of 52.4429 in Interval 3. (Ferdiansyah et al., 2023) achieved a MAPE of 5.31%, whereas our model reported a MAPE of 1.3285% in Interval 3.

3. Comparison with Past Research for Ripple

Table 18 showcases the predictive performance of our models for the next-day closing price prediction of Ripple (XRP). Among the models, the CNN-BiLSTM with a composite of Technical + Fundamental indicators (FI+TI) consistently stands out, especially in Interval 4. In Interval 4, the CNN-BiLSTM model with FI+TI as input achieved an MAE of 0.0054. This is considerably lower than the MAE of 0.0064 reported by (Altan et al., 2019). Furthermore, our model's RMSE of 0.0059 surpasses the RMSE values presented by all the mentioned studies: 0.0088 by (Altan et al., 2019), 0.0123 by (P. Kumar et al., 2022), 0.007 by (Livieris et al., 2021), and notably the 0.0820 by (Ferdiansyah et al., 2023). In terms of the MAPE metric, our model's 0.3937% is significantly better than the 1.47% by (Altan et al., 2019), 0.9042% by (P. Kumar et al., 2022) and 5.37% by (Ferdiansyah et al., 2023)

4. Comparison with Past Research for Litecoin

Table 19 illustrates the predictive accuracy of our models for the next-day closing price prediction of Litecoin (LTC). Among the various models, the CNN-BiLSTM using a combination of Technical and Fundamental indicators (FI+TI) as input variables consistently delivers superior results, particularly in Interval 4. In this interval, the CNN-BiLSTM model achieved an MAE of 0.8602, notably lower than the MAE of 1.1066 reported by (Altan et al., 2019). Moreover, our model's RMSE of 0.1209 outperforms the RMSE values reported by previous studies: 1.7989 by (Altan et al., 2019) and 0.825 by (Hamayel & Owda, 2021). It is also worth highlighting that our model's RMSE is significantly better than the 0.45144 presented by (Tanwar et al., 2021b) for their next-day forecast. In terms of the MAPE metric, our model's performance of 1.2054% is substantially improved compared to the 2.77% by (Altan et al., 2019), even though it is higher than the 0.2116% reported by (Hamayel & Owda, 2021). For the 3-day and 7-day forecasting horizons, the CNN-BiLSTM model using Technical and Fundamental indicators (FI+TI) as input variables continues to demonstrate its efficacy. For the 3-day forecast, our model achieved an RMSE of 0.1603, significantly better than the 0.1450 reported by (Tanwar et al., 2021b). Similarly, for the 7-day forecast, our model's RMSE of 0.8342 is notably superior to the 0.51351 presented by (Tanwar et al., 2021b). This further underscores the robustness and predictive power of the CNN-BiLSTM model across different forecasting horizons.

5. **Comparison with Past Research for DASH**

   Table 20 presents the predictive performance of all models for the next-day closing price prediction of DASH coin. In Interval 4, the CNN-BiLSTM model using Technical and Fundamental indicators (FI+TI) as input variables achieved an MAE of 0.0551, RMSE of 0.1079, and MAPE of 1.0581%. Compared to previous studies, our model's MAE and RMSE are notably superior to the MAE of 2.0746 and RMSE of 2.7776 reported by (Altan et al., 2019) Furthermore, our model's MAE of 0.0551 is also better than the 0.0805 presented by (Parekh et al., 2022). Our model's RMSE is considerably lower than the RMSE of 0.138564 reported by (Patel et al., 2022) for forecasting the next day's price. Additionally, our model's MAPE of 1.0581% is considerably better than the 4.7928% by (Parekh et al., 2022) and is closely aligned with the 1.47% by Altan et al. (2019). For multi-step forecasts spanning 3 and 7 days, Table 25 compares our models' performance for the DASH coin's closing price during Interval 4. When focusing on the 3-day forecast, the CNN-BiLSTM model using Technical and Fundamental indicators (FI+TI) as input variables achieved an RMSE of 0.1338, which is nearly the same as the RMSE of 0.13674 reported by (Patel et al., 2022). However, for the 7-day forecast, our model's RMSE of 0.1223 is notably better than the 0.13341 presented by (Patel et al., 2022).

It is evident that the BiLSTM and CNN-BiLSTM models, especially those utilizing a combination of Technical and Fundamental indicators, consistently outperformed the ANN models in both the 3-day and 7-day forecasting horizons. In particular, the CNN-BiLSTM with FI+TI inputs demonstrated the lowest forecasting errors among all models for these multi-step forecasts, emphasizing its superior predictive capabilities for the DASH coin's closing price over extended periods.

4.2 Objective 2: Price Formation Factors of Bitcoin and Altcoins

Objective 2: *To determine the price formation factors of Bitcoin and Altcoins*

4.2.1 Results for Objective 2

In this section, we summarize the findings of the SHAP analysis conducted on our top-performing model, the CNN-BiLSTM, utilizing a combination of technical and fundamental indicators. This analysis aids in explaining the key factors influencing price formation of Bitcoin and Altcoins.

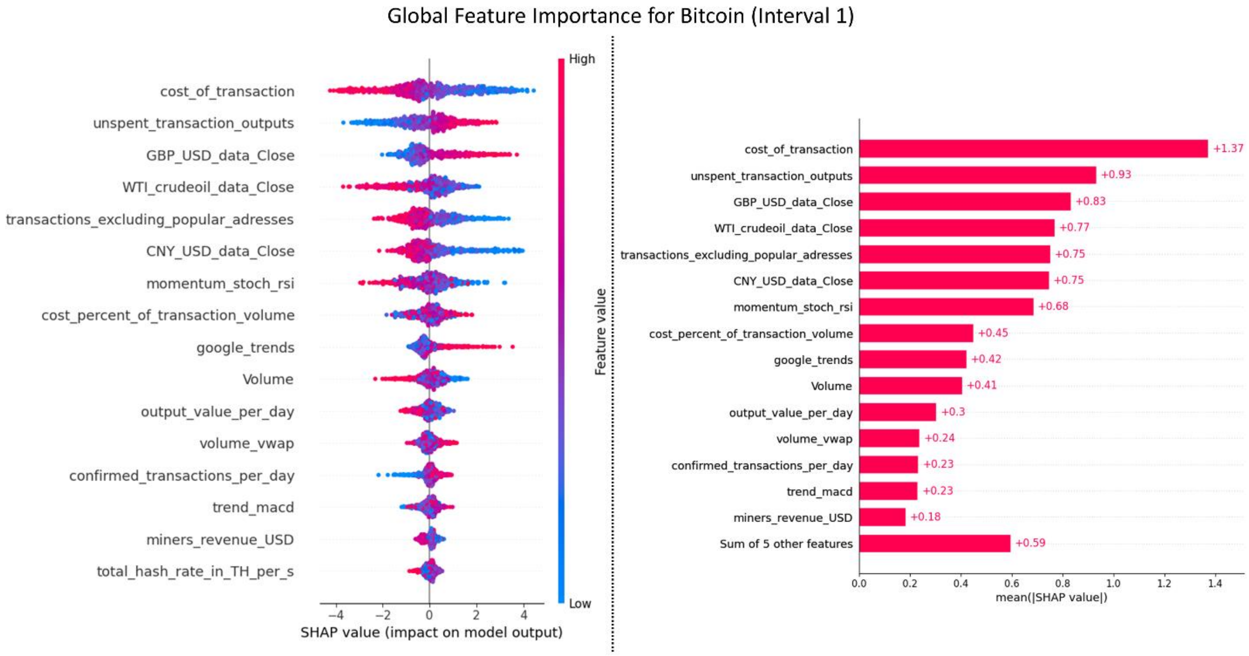


Figure 20. These plots illustrate the main determinants of Bitcoin's price formation. On the left, summary plot derived from SHAP values for the best performing CNN-BiLSTM model with Fundamental and Technical Indicators as the input dataset for Bitcoin's Interval 1 (29-Apr-13 to 31-Jul-16) is displayed. Blue values in the summary plot indicate a negative impact on the model's prediction, while red values signify a positive impact. On the right, a bar plot displays the aggregated magnitude of the SHAP values for each feature.

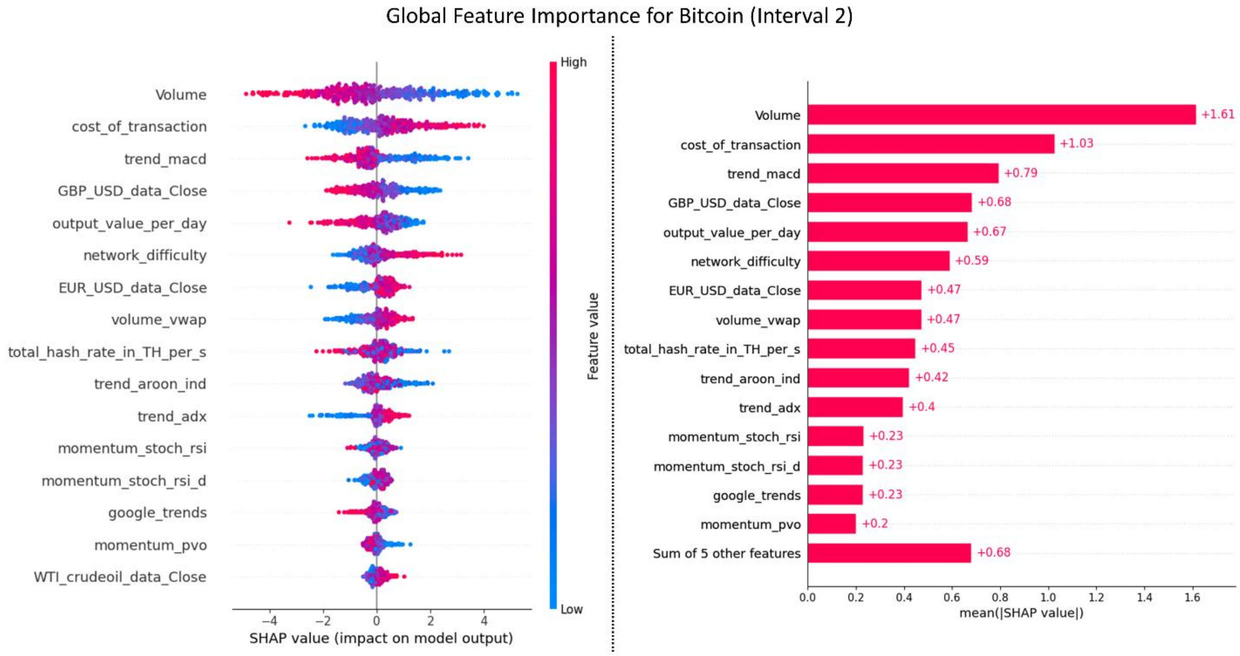


Figure 21. Summary plot and Bar plot obtained from SHAP analysis illustrate the Global Feature Importance of Bitcoin's Interval 2 (29-Apr-13 to 30-Apr-17).

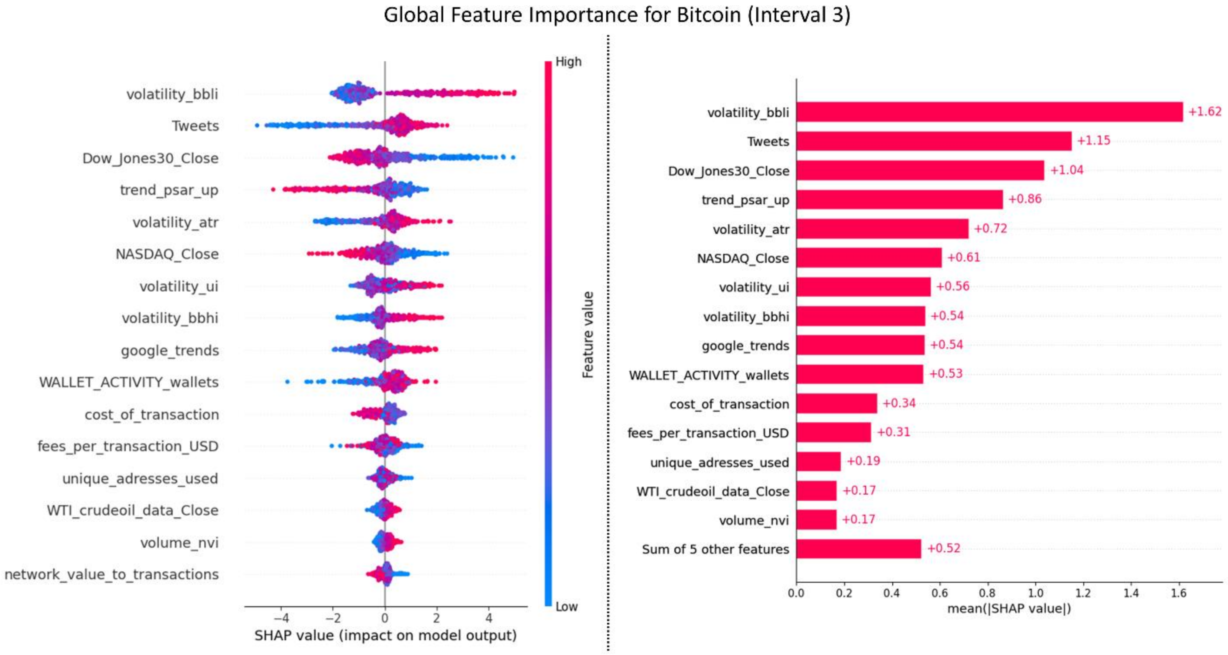


Figure 22. Global Feature Importance for Bitcoin's Interval 3 (29-Apr-13 to 31-Dec-17), showing the shift in feature influence compared to previous intervals.

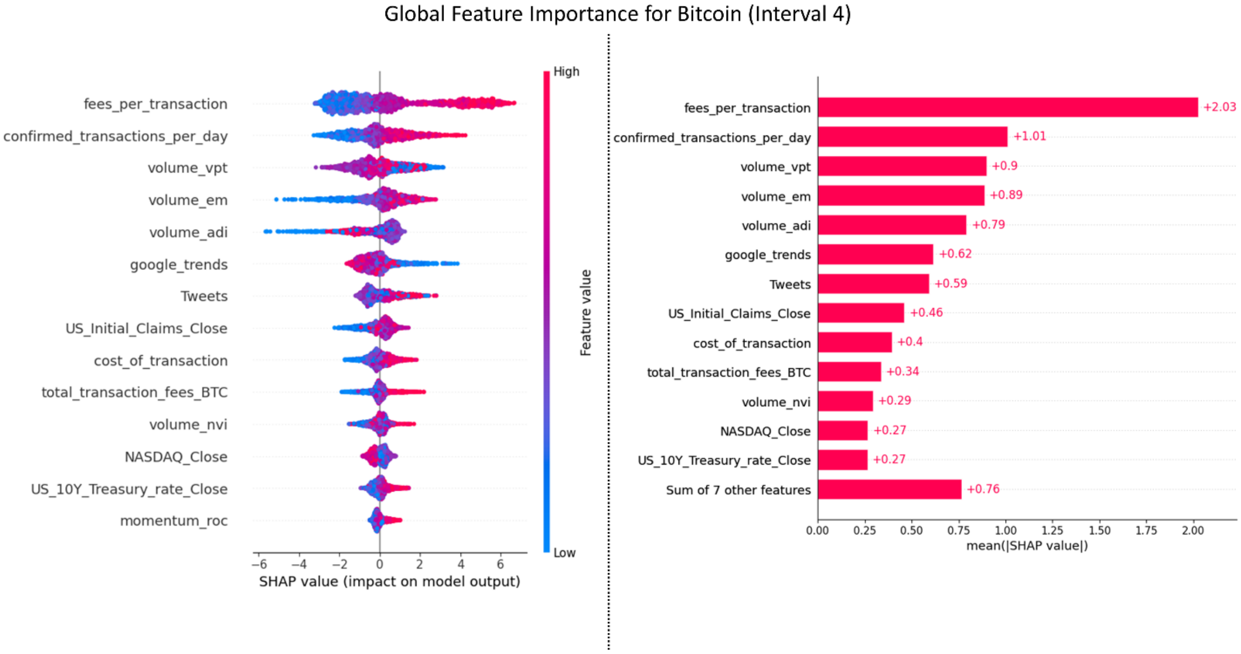


Figure 23. Global Feature Importance summary plot derived from SHAP values for the best performing CNN-BiLSTM model with Fundamental and Technical Indicators as the input dataset for Bitcoin's Interval 4 (2-Jan-17 to 2-Apr-23). Interval 4 is the representative of the most recent data.

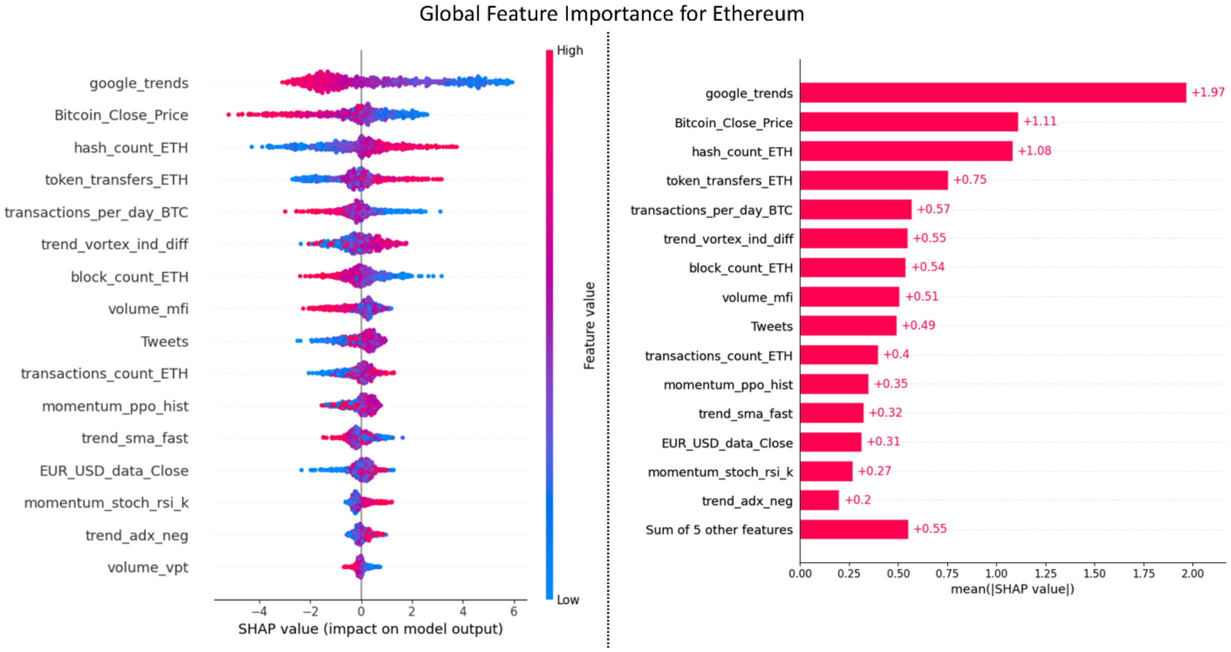


Figure 24. Ethereum's Interval 4 (2-Jan-17 to 2-Apr-23) Feature Importance. The left plot provides a detailed view of feature impact on Ethereum's price, while the right offers an aggregated perspective.

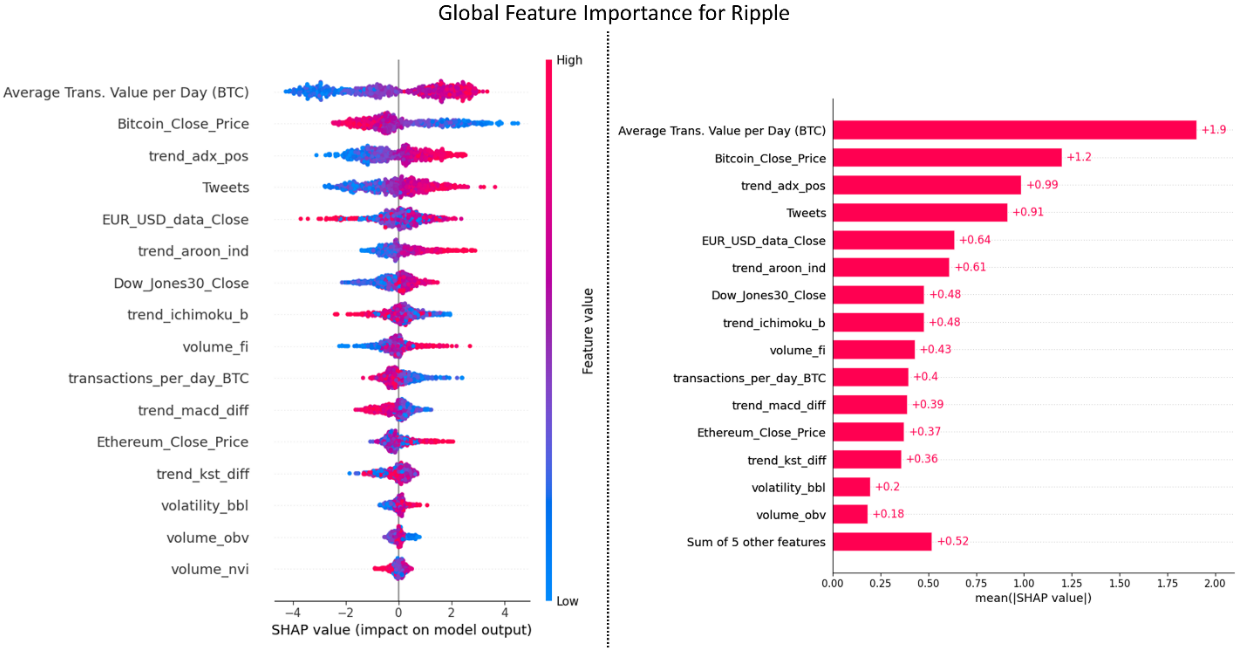


Figure 25. Ripple's Interval 4 (2-Jan-17 to 2-Apr-23) Feature Analysis. The plots highlight the key features driving the price formation factors of Ripple.

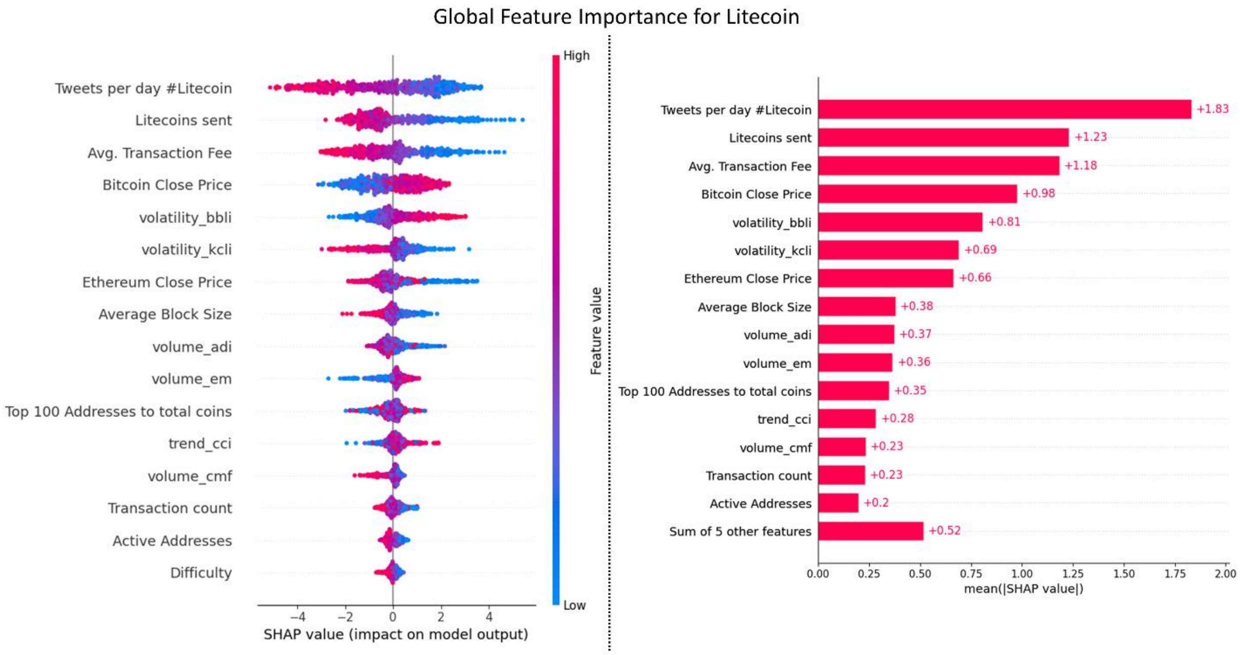

Global Feature Importance for Litecoin

Figure 26. Litecoin's Interval 4 (2-Jan-17 to 2-Apr-23) Feature Breakdown. The juxtaposition of the summary and bar plots offers a comprehensive view of feature influence.

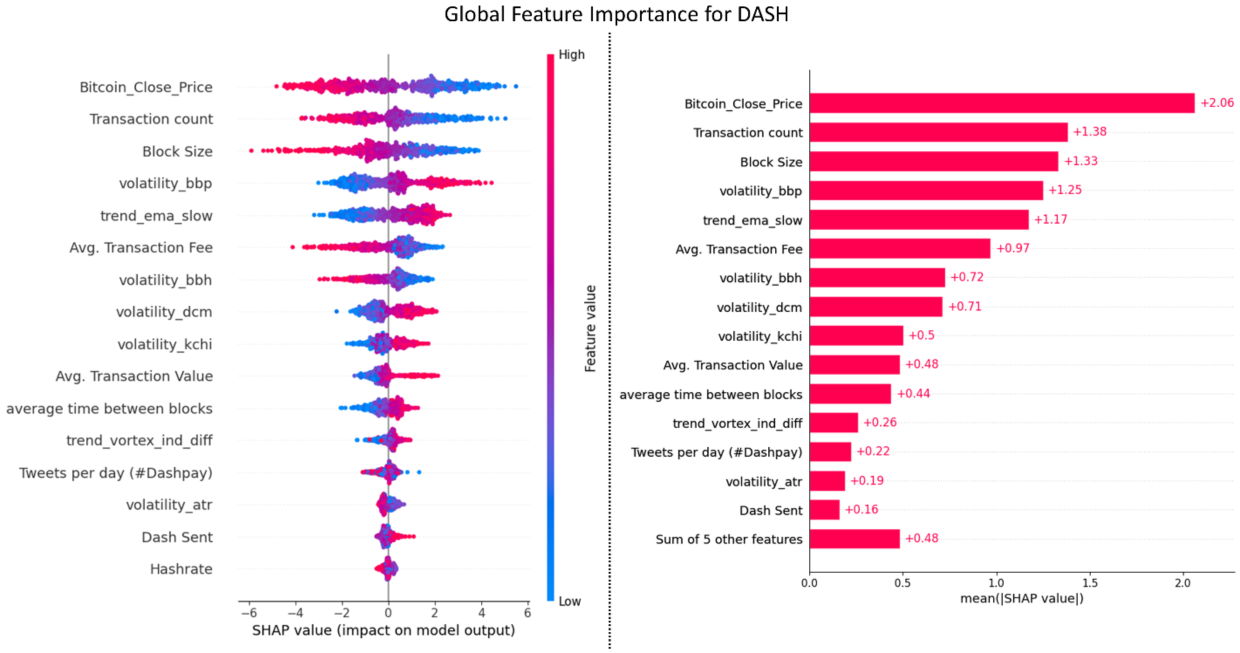

Global Feature Importance for DASH

Figure 27. DASH's Interval 4 (2-Jan-17 to 2-Apr-23) Feature Insights. The plots shows the primary drivers behind the CNN-BiLSTM (FI+TI) model's predictions and for the most recent interval.

In the summary plot, each dot signifies the SHAP value for a single feature for a single observation. The color scheme in this plot provides additional insight into the relationship between feature values and model output. For instance, high values of a feature, represented by red dots, that tend to have positive SHAP values, indicate they push the model’s output higher. Conversely, low feature values, represented by blue dots, that tend to have negative SHAP values, suggest they push the model’s output lower.

4.2.1.1 Interval 1: Bitcoin's Inception and Early Adoption Phase

Figure 20 reveals the global feature importance derived from the SHAP analysis for Interval 1. Interval 1 consists of the period from 29-Apr-13 to 31-Jul-16, corresponding to the initial phases of Bitcoin's development. One notable characteristic of this stage was the significant influence of Blockchain network-related factors. As illustrated in Figure 20, the SHAP analysis identified several Blockchain related factors as important, including the number of unspent Bitcoin transaction outputs, the cost of processing Bitcoin transactions, the number of confirmed Bitcoin transactions per day, and the output value of Bitcoin per day. About eight Blockchain-related features show a dominant influence on Bitcoin's price in this interval. The early market dynamics also show the influence of global currency pairs, specifically GBP to USD and CNY to USD. This is likely due to the significant adoption of Blockchain in Europe and China during these early days. Google trends also indicate a heightened interest in Bitcoin during this period. However, there is limited influence from external macroeconomic factors during this period, apart from crude oil and global currencies to US dollar ratios. This importance of Blockchain-related explanatory variables, alongside global currency to US dollar ratios, aligns with the findings presented by (W. Chen et al., 2021) for a comparable overlapping timeframe.

4.2.1.2 Interval 2: Expansion Phase and Shift to Market-Related Factors

Figure 21 displays the SHAP summary plot for Interval 2, which spans from 29-Apr-13 to 30-Apr-17 marked an expansion phase characterized by heightened public attention, an influx of traders, speculators, and a broader spectrum of market participants primarily seeking short-term gains. Particularly, the Blockchain-centric factors that dominated Interval 1 receded in their influence during this phase. Instead, market-related factors such as transaction volume, cost of transaction, and global currency ratios, specifically the British Pound Sterling to US Dollar (GBP to USD) and the Euro to US Dollar (EUR to USD), emerged as significant price formation factors of Bitcoin.

Furthermore, the emergence of technical indicators such as Moving Average Convergence Divergence (trend_macd), Volume Weighted Average Price (volume_vwap), Average Directional Index (trend_adx), Percentage Volume Oscillator (momentum_pvo), and Aroon Indicator (trend_aroon_ind) as influential factors emphasizes the important role of technical analysis in this interval. The heightened importance of these technical indicators suggests a shift in trading strategies, with market participants increasingly relying on price patterns and market trends for short-term trading decisions.

4.2.1.3 Interval 3: Transitional Phase during Initial Market Peak

Figure 22 depicts the SHAP summary plot for Bitcoin's Interval 3, which spans from 29-Apr-13 to 31-Dec-17. During this time, Bitcoin reached its then-all-time maximum price of $16,000, attracting significant attention from various financial market participants, particularly from May to December 2017. The volatility index (Bollinger band indicator - bbli) is the predominant factor affecting Bitcoin pricing, as denoted by its highest mean SHAP value of +1.62

This elevated SHAP value indicates that fluctuations in volatility are highly indicative of changes in Bitcoin's price. Additionally, social media dynamics, proxied by the volume of Tweets, emerges as a significant predictor, underscoring the impact of digital social constructs on cryptocurrency valuation. The closing values of entrenched financial indices, specifically the Dow Jones and NASDAQ, are also observed to be influential, signaling an interconnection between Bitcoin and traditional equity markets. The pertinence of public interest and market sentiment is further corroborated by the positive association with Google search trends. Wallet activity metrics and the economic factors of Bitcoin, such as transaction costs and fees, are also notable contributors, albeit with less pronounced SHAP values. Interestingly, the aggregate influence of the five least impactful features yields a cumulative SHAP value of +0.52, illustrating their collective, yet comparatively moderate, effect on the model's predictions. The scatter of SHAP values depicted in the figure illustrates the heterogeneity in the impact of each feature, offering insights into the complex and diverse nature of factors influencing Bitcoin's pricing behavior within the studied interval.

The importance of features such as NASDAQ, public attention factors (Google Trends, and Tweets), transaction fees, and transaction volume-related metrics aligns with the findings of (W. Chen et al., 2021) during a similar investigation period.

4.2.1.4 Interval 4: Price Formation Through Multiple Market Cycles

Figure 23 presents the SHAP summary plot for Bitcoin's Interval 4, from 2-Jan-17 to 2-Apr-23. This period is important because it encompasses the most recent market data and multiple bull-bear cycles including the all-time high prices observed in Dec 2017, April, 2021 and Nov 2021 and subsequent market corrections, reflecting the maturation and increased volatility of the Bitcoin market. The most important factor in this interval is 'fees_per_transaction' (Fees per Transaction in USD). The importance of this factor suggests that in the operational efficiency of the Bitcoin network, the transaction costs remain a decisive factor. High fees discourage users, reducing demand, while low fees encourage increased transactions and demand, directly influencing the price balance. The significance of this feature aligns with the observations made by (W. Chen et al., 2021) who reported a similar finding. The second most important factor is the 'confirmed_transactions_per_day' (Number of Confirmed Transactions per Day), reinforcing that Blockchain network activity, representing both demand and supply-side dynamics, is integral to Bitcoin's price determination.

The subsequent influential factors, 'volume_vpt' (Volume Price Trend Indicator), 'volume_em' (Ease of Movement Value), and 'volume_adi' (Accumulation/Distribution Index), are volume-centric indicators. Their significance in this interval emphasizes the key role of market liquidity, capital flows, and trading volumes in shaping Bitcoin's price. The prominent influence of 'google_trends' (Google Search Trends for Bitcoin) and 'Tweets' (Number of Bitcoin-related Tweets) in this interval restates the importance of public sentiment, awareness, and digital discourse in driving demand and influencing Bitcoin's price. These findings corroborate with the findings of (García-Medina & Huynh, 2021). Lastly, the inclusion of broader macroeconomic indicators such as 'US_Initial_Claims_Close' (US Initial Unemployment Claims), 'NASDAQ_Close' (NASDAQ Composite Index Closing Value), and 'US_10Y_Treasury_rate_Close' (US 10-Year Treasury Rate Closing Value), coupled with market momentum indicators like 'momentum_roc' (Rate of Change Momentum Indicator), indicates that Bitcoin's price dynamics in Interval 4 are a product of a multifaceted set of factors.

While our research is extensive, employing a comprehensive dataset not previously explored in existing literature, it is essential to approach comparisons with prior studies cautiously due to their more limited scope. Specifically, (W. Chen et al., 2021) identified transaction volume, tweets, and transaction fees as important determinants of Bitcoin's price during a similar period. Transaction fees, volume, Google trends and volume related metrics were also found significant in (Kukacka & Kristoufek, 2023). (Corbet et al., 2020) highlighted the importance of Unemployment (US Initial Unemployment Claims) in their study which is consistent with our findings. Furthermore, (García-Medina & Huynh, 2021) highlighted the significance of Google trends. However, these studies did not incorporate technical indicators in their analyses and utilized a restricted number of fundamental indicators.

4.2.1.5 Ethereum's Price Formation Factors

Figure 24 presents the SHAP summary plot for Ethereum's price formation factors for the Interval 4. This period is crucial as it captures the most recent market data, reflecting the maturation and heightened volatility of the Ethereum market. Ethereum, like Bitcoin, experienced significant price fluctuations during this period, including reaching peak prices and undergoing market corrections. The most important price formation factor is 'google_trends' (Google Search Trends for Ethereum), emphasizing the role of public sentiment and awareness in Ethereum's price dynamics. The 'Bitcoin close price' is the next most significant factor, indicating that Bitcoin strongly influences the price formation of Ethereum.

Ethereum's blockchain-specific factors such as 'hash_count_eth' (Hash Count of Ethereum), 'token_transfers_eth' (Token Transfers on Ethereum), and 'block_count_eth' (Block Count of Ethereum) highlight the importance of Ethereum's network activity in its price determination.

These blockchain-specific factors provide information on the level of activity and usage of the Ethereum network, which can directly impact its price. For example, a high number of token transfers on Ethereum indicates increased demand and usage of the network, which can drive up the price of Ethereum. Similarly, a higher block count suggests a higher level of network activity and can also contribute to price volatility.

The 'transactions_per_day_btc' (Number of Bitcoin Transactions per Day) factor underscores the interconnectedness of the cryptocurrency market, where the activity in Bitcoin can influence Ethereum's price dynamics. Technical indicators, including 'trend_vortex_ind_diff' (Vortex Indicator Difference), 'volume_mfi' (Money Flow Index), 'momentum_ppo_hist' (Percentage Price Oscillator Histogram), 'trend_sma_fast' (Simple Moving Average - Fast), and 'volume_vpt' (Volume Price Trend Indicator), are significant in this interval. Their prominence suggests that like Bitcoin, traders, and investors in the Ethereum market also rely on technical analysis for their trading and investment actions. The influence of 'Tweets' (Number of Ethereum-related Tweets) in this interval highlights the role of social media sentiment in Ethereum's price, reflecting the community's sentiment and potential market-moving events. Furthermore, the 'EUR_USD' (Euro to US Dollar exchange rate) factor points to the influence of macroeconomic elements, suggesting that global financial dynamics and currency movements can have repercussions on Ethereum's valuation.

Our findings are consistent with those of (Angela & Sun, 2020), who discovered that Bitcoin prices substantially affect Ethereum prices, and (H. M. Kim et al., 2021), who discovered that Blockchain information of the Ethereum network plays a significant role in determining Ethereum prices.

4.2.1.6 Ripple’s Price Formation Factors

Figure 25 presents the SHAP summary plot for Ripple's price formation factors for Interval 4. Like Bitcoin and Ethereum, Ripple also witnessed significant price fluctuations during this time, achieving peak prices and undergoing market corrections.

The primary driver of Ripple's price in this interval is the Average Transaction Value per Day (BTC), emphasizing the influence of Bitcoin transaction values on Ripple's price dynamics. The Bitcoin and Ethereum Close Prices are also important factors, indicating that both Bitcoin and Ethereum, as leading cryptocurrencies, substantially impact Ripple's price formation. Our findings corroborate with the results of (Agyei et al., 2022) who showed a high co-movement of Ripple prices with that of Bitcoin and Ethereum prices.

It can also be observed in Figure 25 that technical indicators also have a strong influence Ripple's price. The Aroon Indicator (trend_aroon_ind) helps identify when Ripple's price is in a trend and the magnitude of that trend. In the SHAP plot, the small blue section towards the negative x-axis suggests that when the Aroon value is low, there is a marginal upward pressure on Ripple's price, indicating a potential bullish trend. Conversely, the extended red section towards the positive x-axis indicates that higher Aroon values exert a more pronounced downward pressure on the coin's price, signaling a bearish trend. This pattern in the SHAP analysis highlights the Aroon Indicator's significance in Ripple's price formation, with market participants likely using it to detect trend changes and their strength.

The influence of number of Ripple-related Tweets (Tweets) which is the next most important factor, highlights the role of social media sentiment in Ripple's price, reflecting the community's sentiment and potential market-moving events. The Global currency ratio Euro to US Dollar exchange rate (EUR_USD), and the Dow Jones 30 Industrial Average point to the influence of macroeconomic elements and global financial market dynamics on Ripple's valuation.

The Positive Directional Movement Index (trend_adx_pos) which is the next most important factor quantifies the strength of an upward trend in Ripple coin's price. It can be seen in the SHAP plot, that the left side of this indicator predominantly appears blue, suggesting a negative SHAP value, while the right side is mostly red, indicating a positive SHAP value. This distribution implies that when the trend_adx_pos value is lower (indicating weaker upward momentum), it tends to exert downward pressure on Ripple's price. Further, Dow Jones 30 is another macro-economic variable that can be seen to influence the Ripple coin’s price.

The Ichimoku Base Line (trend_ichimoku_b) represents the midpoint of the last 26 price bars, serving as a dynamic support or resistance for Ripple's price. A thin red line on the left of the SHAP plot indicates that when Ripple's price drops below this threshold, there is a minor upward pressure on the coin's price. In contrast, when the price is substantially above the baseline, a prominent blue region at axis 0 indicates a stronger downward pressure. This pattern demonstrates the significance of the baseline in Ripple's price formation, with traders likely adjusting their strategies accordingly.

4.2.1.7 Litecoin’s Price Formation Factors

Figure 26 presents the SHAP summary plot for Litecoin's price formation factors during Interval 4. Like other cryptocurrencies, the price trajectory of Litecoin is significantly influenced by social media sentiment, Bitcoin and Ethereum's prices, Litecoin's Blockchain variables, and a few technical indicators. The SHAP analysis indicates that the number of daily Tweets containing the #Litecoin hashtag is the most significant factor in determining the price of Litecoin. Following this, the variable 'Litecoins sent' represents the volume of Litecoin transferred on the network. A higher volume can indicate increased adoption, usage, and, potentially, heightened demand, leading to price appreciation. The Average Transaction Fee (in

LTC) is another crucial feature, representing the cost of executing transactions on the Litecoin network. An increase in this fee might suggest network congestion or a surge in demand for transaction processing. The impact of other prominent cryptocurrencies is visible, as indicated by the presence of Bitcoin Close Price (Bitcoin_Close_Price) and Ethereum Close Price (Ethereum_Close_Price) as important factors. Litecoin blockchain network variables such as Average Block Size (Average_Block_Size), Transaction count (Transaction_count), Active Addresses (Active_Addresses), and Difficulty (Difficulty) are also important in Litecoin's price formation.

The Average Block Size indicates the amount of data processed in each Litecoin block, with larger sizes suggesting robust network activity. Transaction count reflects the number of transactions processed, while Active Addresses denotes the number of unique addresses participating in transactions on a given day. The difficulty, representing the computational challenge of mining a new block, indicates the security and miner commitment to the Litecoin network. Technical indicators, including Bollinger Bands Lower Indicator (volatility_bbli), Keltner Channel Lower Indicator (volatility_kcli), Accumulation/Distribution Index (volume_adi), Ease of Movement Value (volume_em), Commodity Channel Index (trend_cci), and Chaikin Money Flow (volume_cmf), play a significant role in this interval. The two volatility indicators, Bollinger Bands Lower Indicator and Keltner Channel Lower Indicator suggest that Litecoin's price is subject to notable fluctuations, reflecting its market volatility.

Our findings are consistent with those of (Angela & Sun, 2020), who demonstrated a significant spillover between Litecoin prices and the prices of Bitcoin and Ethereum.

4.2.1.8 DASH Coin’s Price Formation Factors

Figure 27 presents the SHAP summary plot for DASH's price formation factors during Interval 4. The price trajectory of DASH, like other cryptocurrencies, is significantly influenced by Bitcoin's price movements, inherent blockchain variables of DASH, and a range of technical indicators.

The Bitcoin Close Price is the most important feature in DASH's price formation. The Transaction count derived from DASH blockchain network is the next most important feature. It reflects the number of transactions processed on the DASH network. A higher transaction count can signify increased adoption and usage, potentially leading to higher demand and price appreciation. The next important feature, Block Size, indicates the amount of data processed in each DASH blockchain block. Larger block sizes suggest more complex or numerous transactions, indicating robust network activity and vice versa. Like other cryptocurrencies, the Average Transaction Fee (in DASH) or the cost associated with executing transactions on the DASH network is another important feature of the price formation. An increasing average fee might indicate network congestion or higher demand for transaction processing, influencing price dynamics. Average Transaction Value (in DASH) is another important feature representing the average value of transactions on the DASH network. A higher average value can suggest larger transfers or investments, potentially indicating significant financial activity or institutional involvement. The average time between blocks (in minutes) is another important factor. Shorter times indicate efficient network processing and higher miner activity, while longer times suggest network congestion or reduced miner participation. Hash rate (measure of computational power) is also an important factor that denotes the total computational power used to mine and process transactions on the DASH network. A rising hash rate signifies increased security and miner commitment to the network, which can be a positive signal for potential investors.

Figure 27 further illustrates the significance of volatility-based technical indicators in the price formulation of DASH coin. Specifically, the Volatility Bollinger Bands Percentage (Volatility_bbp) measures DASH's price in relation to its upper and lower bands, offering insights into potential liquidity constraints and market sentiment. Similarly, the Volatility Bollinger Bands High (volatility_bbh) represents the upper threshold of price, indicating potential bullish sentiment and possible liquidity limitations when the price consistently approaches this band. The Volatility Keltner Channel High (volatility_kchi) uses the Average True Range to determine overextended price levels, suggesting strong bullish sentiment when the price frequently interacts with the upper channel. Lastly, the Volatility Donchian Channel Middle (volatility_dcm) acts as a sentiment pivot, with prices above and below this line indicating bullish and bearish sentiments, respectively.

The influence of 'Tweets per day (#Dashpay)' in this interval highlights the role of social media sentiment in DASH's price, reflecting the community's sentiment and potential market-moving events.

4.2.2 Main Contributions and Novelty

1. **Explainable AI for Deep Learning:** The main novelty of this study (for this objective) is the use of Explainable AI to provide explanations for a Deep Learning framework that determines the price formation factors of Bitcoin and Altcoins. This is a significant advancement as most previous studies primarily focused on Bitcoin, leaving the dynamics of Altcoins relatively unexplored.
2. **Comprehensive Dataset:** In contrast to previous studies that primarily relied on either fundamental indicators, technical indicators, or univariate forecasting, our research utilizes a dataset that is not only more comprehensive but also more representative of the underlying factors of cryptocurrency prices. This dataset encompasses both fundamental and technical indicators, offering a more complete understanding of the factors that influence cryptocurrency prices.

3. **Inclusion of 85+ Technical Indicators:** Unlike previous studies that limited their analysis to commonly used technical indicators, our research incorporates over 85 technical indicators. For context, previous studies often relied on indicators like Moving Averages, Relative Strength Index (RSI), Bollinger Bands, and MACD (Moving Average Convergence Divergence). Our approach, by contrast, provides a richer and more detailed insight into price dynamics.
4. **Segmented Analysis:** Our study adopts a segmented approach, dividing the analysis into distinct intervals. This allows us to capture the evolving nature of cryptocurrency markets and understand how different factors gain or lose prominence during various market phases. Specifically, we have identified price formation factors for both Bitcoins and Altcoins across different market conditions.

Additionally, for this objective, we juxtapose our findings with prior research, allowing for validation of results and identification of novel insights. This comparative approach helped in understanding the consistency and variability in price determinants over time.

4.3 Results for Objective 3: Factors for Price Inconsistencies Across Exchanges

0bjective 3*:* *To identify the factors responsible for price inconsistencies across different major cryptocurrency exchanges*

4.3.1 Descriptive Statistics of Prices and Price Inconsistencies

In this objective, we focus on five cryptocurrencies: Bitcoin, Ethereum, Ripple, Litecoin, and DASH, comparing their prices across three major cryptocurrency exchanges: Binance (Global), Kraken, and Coinbase Pro. Our primary aim is to identify the main factors responsible for inconsistencies across these markets. To achieve this, we first compare the descriptive statistics of the prices of each cryptocurrency across these exchanges to identify price inconsistencies. We then transition to metrics of price discrepancies, both absolute and relative terms. The relative price discrepancy offers a perspective on the variation in terms of the general price level, which is crucial for volatile entities like cryptocurrencies. We then present the findings from our price inconsistencies model, including its hyperparameters and performance metrics such as RMSE, MAE, and MAPE. Subsequently, we use the SHAP analysis to present this objective's main results and findings, where we analyze the key determinants and their contribution to price inconsistencies within the studied exchanges.

Tables 26 to 30 display the descriptive statistics of five major cryptocurrencies: Bitcoin, Ethereum, Ripple, Litecoin, and DASH across Binance, Coinbase Pro, and Kraken.

Table 26. Descriptive statistics for closing prices of Bitcoin (BTC)

Statistic	Binance	Coinbase Pro	Kraken
Mean	17870.30	18270.05	17833.10
Median	10131.52	10109.46	10118.36
Mode	1179.97	5575.83	24178.96
Standard Deviation	16427.45	16061.51	16431.97
Kurtosis	0.24	0.34	0.27
Skewness	1.18	1.24	1.19
Minimum	777.76	3232.51	784.28
Maximum	67566.83	67549.14	67617.02
Count	2291	2291	2291

Source: Author's Calculations**.**

Table 27. Descriptive statistics for closing prices of Ethereum (ETH)

	Binance	Coinbase Pro	Kraken
Mean	1040.86	1039.11	1019.41
Median	408.14	407.85	408.27
Mode	320.88	313.54	1678.92
Standard Deviation	1117.99	1119.28	1134.80
Kurtosis	0.92	0.93	0.87
Skewness	1.37	1.37	1.35
Minimum	84.31	84.12	8.06
Maximum	4812.09	4810.97	4815.01
Count	2291	2291	2291

Source: Author's Calculations**.**

Table 28. Descriptive statistics for closing prices of Ripple (XRP)

	Binance	Coinbase Pro	Kraken
Mean	0.48	0.47	0.47
Median	0.35	0.35	0.35
Mode	0.22	0.22	0.37
Standard Deviation	0.35	0.33	0.36
Kurtosis	10.80	6.83	9.44
Skewness	2.56	2.18	2.31
Minimum	0.14	0.14	0.01
Maximum	3.38	2.78	3.40
Count	2291	2291	2291

Source: Author's Calculations**.**

Table 29. Descriptive statistics for closing prices of Litecoin (LTC)

	Binance	Coinbase Pro	Kraken
Mean	91.16	95.08	91.08
Median	67.96	67.29	67.94
Mode	90.68	60.55	81.70
Standard Deviation	63.04	59.17	63.20
Kurtosis	1.80	2.39	1.82
Skewness	1.33	1.56	1.34
Minimum	3.71	23.12	3.73
Maximum	386.45	388.28	384.67
Count	2291	2291	2291

Table 30. Descriptive statistics for closing prices of DASH

	Binance	Coinbase Pro	Kraken
Mean	189.39	184.31	166.64
Median	122.04	121.81	104.32
Mode	326.01	296.19	56.42
Standard Deviation	184.21	177.48	182.60
Kurtosis	10.96	10.74	13.62
Skewness	2.82	2.84	3.31
Minimum	32.26	32.11	11.21
Maximum	1550.85	1437.46	1493.59
Count	2291	2291	2291

Source: Author's Calculations**.**

In Tables 26 to 30, we observe distinct price patterns and inconsistencies across cryptocurrencies and exchanges. These variations arise from the unique market conditions, trading behaviors, and external factors inherent to each cryptocurrency and exchange pairing. It can be observed that there are variations in the mean prices of cryptocurrencies across crypto exchanges. These differences highlight that the same cryptocurrency does not adhere to the Law of One Price (LOOP) when traded on different exchanges. The observations are further explored in the analysis below:

4.3.1.1 Price Discrepancy Observations Across Exchanges

1. **Variability in Mean Prices:** Different mean prices for cryptocurrencies across the exchanges highlight price inconsistencies. The mean closing prices are closely clustered for Bitcoin across the exchanges, though Coinbase Pro holds a slightly higher mean compared to the other two. Bitcoin's mean price on Coinbase Pro is 18270.05, while on Kraken it is 17833.10 and on Binance it is 17870.30.

2. **Significant Standard Deviations**: The standard deviations for each cryptocurrency are relatively high on all exchanges, indicating a high level of price volatility. Kraken and Binance display almost identical standard deviations, indicating a similar volatility pattern for Bitcoin on these platforms. For example, Bitcoin's standard deviation on Binance is 16427.45, on Coinbase Pro it is 16061.51, and on Kraken it is 16431.97.

3. **Presence of Skewness**: A non-zero skewness value across the cryptocurrencies suggests that the returns distribution is not symmetrical. Positive skewness values hint at the presence of larger values on the right side of the mean, suggesting potential outliers or extreme values impacting the prices. Ripple, for instance, has a skewness of 2.56 on Binance, 2.18 on Coinbase Pro, and 2.31 on Kraken.

4.3.1.2 Exchange Specific Analysis

1. **Coinbase Pro**: This platform displays an elevated mean price for cryptocurrencies, particularly Bitcoin and Litecoin, hinting at a distinct user demographic or trading approach. For example, Bitcoin's mean price on Coinbase Pro is 18270.05, which is higher than Binance's 17870.30 and Kraken's 17833.10.
2. **Kraken**: For Ethereum and DASH, Kraken exhibits a slightly lower mean compared to the other exchanges. This could be attributed to specific market dynamics or trading behaviors associated with Kraken's user base. It can be observed that the mean for Ethereum and DASH is marginally reduced, with Ethereum's mean price on Kraken being 1019.41, lower than Binance's 1040.86 and Coinbase Pro's 1039.11.
3. **Binance**: Binance's Kurtosis values, especially for Ripple, are substantially high, indicating the presence of heavy tails or outliers in the distribution and point to the presence of extreme price events. Specifically, the kurtosis value for Ripple on Binance is 10.80, which is higher than Coinbase Pro's 6.83.

4.3.1.3 Cryptocurrency Specific Analysis

1. **Bitcoin (BTC)**:
    i. The mean closing prices are closely clustered for Bitcoin across the exchanges, though Coinbase Pro holds a slightly higher mean compared to the other two.
    ii. Kraken and Binance display almost identical standard deviations, indicating a similar volatility pattern for Bitcoin on these platforms.
    iii. The mode for Kraken stands out as substantially higher, suggesting a particular price point at which Bitcoin was frequently traded.
2. **Ethereum (ETH)**:
    i. Ethereum's mean prices are relatively close across all exchanges, but Kraken is slightly lower, hinting at different trading conditions or behaviors.
    ii. The minimum price for Ethereum on Kraken is notably lower than on the other exchanges, pointing to a potential outlier or an extreme price drop at some given point.
3. **Ripple (XRP)**:
    i. The kurtosis values for Ripple are significantly higher, especially on Binance, indicating the presence of extreme values or outliers in the price distribution for Ripple.
    ii. Kraken's maximum price stands out, hinting at price spikes on this platform.

4. **Litecoin (LTC)**:
    i. Coinbase Pro has a higher mean price for Litecoin, which could indicate a different demand-supply dynamic or trading behavior for Litecoin on this exchange.

    ii. The minimum price of Litecoin on Binance and Kraken are close, but it is substantially higher on Coinbase Pro, suggesting that Coinbase Pro might have been less affected by any drastic price drops that impacted the other exchanges.

5. **DASH**:
    i. Kraken's mean price for DASH is significantly lower than both Binance and Coinbase Pro, indicating a different trading dynamic for DASH on Kraken.
    ii. The kurtosis values are high across all exchanges, but notably higher for Kraken, pointing to the presence of extreme values in the price distribution for DASH on this platform.

After analyzing the descriptive statistics of closing prices, we quantify the degree of price inconsistency by examining the absolute and relative price discrepancies of cryptocurrencies across exchanges. The absolute price discrepancy represents the direct difference between the highest and lowest prices. Conversely, the relative price discrepancy describes this difference as a percentage of the cryptocurrency's average price, providing a clearer perspective on the price discrepancy in terms of its overall price level. Tables 31-35 display the descriptive statistics for these price discrepancies.

4.3.2 Absolute and Relative Price Discrepancies of Each Exchange

Table 31. Descriptive Statistics of Absolute and Relative Price Discrepancies (referred to as "Spread" in this and subsequent tables) Across Exchanges for Bitcoin (BTC). This table compares the magnitude of price discrepancies between exchanges, using both absolute and relative measures, to assess the consistency in BTC pricing across platforms.

	Absolute Spread	**Relative Spread**
Mean	65979.48	3.66
Standard Error	831.52	0.07
Median	66789.07	3.73
Standard Deviation	1440.24	0.13
Kurtosis	-1.50	-1.50
Skewness	-0.71	-0.70
Range	2516.11	0.23
Minimum	64316.63	3.52
Maximum	66832.74	3.74
Count	3	3

Source: Author’s Calculations**.**

Table 32. Descriptive Statistics of Absolute and Relative Price Discrepancies Across Exchanges for Ethereum (ETH).

	Absolute Spread	**Relative Spread**
Mean	4753.86	4.59
Standard Error	26.54	0.06
Median	4727.78	4.54
Standard Deviation	45.97	0.10

Kurtosis	-1.50	-1.50
Skewness	0.71	0.70
Range	80.09	0.17
Minimum	4726.85	4.54
Maximum	4806.94	4.71
Count	3	3

Source: Author’s Calculations**.**

Table 33. Descriptive Statistics of Absolute and Relative Price Discrepancies Across Exchanges for Ripple (XRP).

	Absolute Spread	**Relative Spread**
Mean	3.09	6.53
Standard Error	0.23	0.50
Median	3.24	6.77
Standard Deviation	0.40	0.87
Kurtosis	-1.50	-1.50
Skewness	-0.59	-0.46
Range	0.75	1.69
Minimum	2.64	5.56
Maximum	3.39	7.26
Count	3	3

Source: Author’s Calculations**.**

Table 34. Descriptive Statistics of Absolute and Relative Price Discrepancies Across Exchanges for Litecoin (LTC).

	Absolute Spread	**Relative Spread**
Mean	376.28	4.07
Standard Error	5.58	0.12
Median	380.94	4.18
Standard Deviation	9.67	0.20
Kurtosis	-1.50	-1.50
Skewness	-0.68	-0.70
Range	17.58	0.36
Minimum	365.16	3.84
Maximum	382.74	4.19
Count	3	3

Source: Author’s Calculations**.**

Table 35. Descriptive Statistics of Absolute and Relative Price Discrepancies Across Exchanges for DASH.

	Absolute Spread	**Relative Spread**
Mean	1468.78	8.15
Standard Error	33.39	0.37
Median	1482.39	7.99
Standard Deviation	57.84	0.65
Kurtosis	-1.50	-1.50

Skewness	-0.41	0.43
Range	113.24	1.26
Minimum	1405.35	7.60
Maximum	1518.59	8.87
Count	3	3

Source: Author's Calculations.

Tables 31-35 present the observations on price discrepancies of cryptocurrencies across various exchanges. From these tables, it is evident that cryptocurrencies exhibit varying degrees of price discrepancy across exchanges.

1. **Magnitude of Bitcoin's Price Discrepancy**: Bitcoin showcases substantial price discrepancies, both in absolute and relative terms, across exchanges. The average absolute discrepancy for Bitcoin is 65979.48, with a range of 2516.11. Its relative discrepancy averages at 3.66%, with a minimal standard deviation of 0.13%.
2. **Elevated Discrepancy for Ethereum**: Ethereum, while having a smaller absolute discrepancy than Bitcoin, manifests a slightly higher relative discrepancy. Ethereum's absolute discrepancy averages at 4753.86 with a range of 80.09, while its relative discrepancy averages at 4.59%, with a standard deviation of 0.10%.
3. **High Relative Discrepancy for Ripple (XRP)**: XRP's absolute discrepancy is modest, yet its relative discrepancy is considerably high, suggesting significant percentage variations in its prices. XRP's absolute discrepancy averages at 3.09 with a range of 0.75. Its relative discrepancy stands at a substantial average of 6.53%.
4. **Litecoin's Price Variability:** Litecoin displays consistent discrepancies in terms of both absolute and relative measures. The average absolute discrepancy for Litecoin is 376.28 with a range of 17.58. Its relative discrepancy averages at 4.07%, with a standard deviation of 0.20%.
5. **DASH's Pronounced Discrepancies**: DASH records a significant relative discrepancy, suggesting substantial percentage variations in its prices across exchanges. DASH's absolute discrepancy stands at 1468.78, with a range of 113.24. Its relative discrepancy is at a notable average of 8.15%.
6. **Uniformity in Kurtosis Across Cryptocurrencies**: All the cryptocurrencies display a kurtosis of -1.50, indicating a consistency in their price discrepancy distributions. This kurtosis value suggests that the distribution of discrepancies is less prone to extreme values than a standard normal distribution.
7. **Directional Variability in Discrepancy Distribution**: The direction of asymmetry, or skewness, varies among the cryptocurrencies. For example, Bitcoin, Litecoin, and DASH exhibit negative skewness, hinting at a more extended left tail in their distributions. In contrast, Ethereum and XRP manifest positive skewness, indicating the opposite effect on their distribution

We can see that the distribution of these discrepancies is mostly consistent without many extreme outliers. In the context of a deep learning model designed to identify factors responsible for price inconsistencies, this consistent distribution suggests that the model can be trained more efficiently without the influence of extreme values. The uniformity in the distribution further indicates that the factors contributing to price inconsistencies are likely systematic or recurrent, enabling the model to more effectively pinpoint and understand these factors.

4.3.3 Model Performance Metrics and SHAP Analysis Results

The initial step in understanding price inconsistencies is to examine the descriptive statistics. After this analysis, we trained our model as described in Algorithm 4 to forecast the price inconsistencies. Table 36 displays the important features obtained from the three-step feature selection process detailed in Algorithm 4. Furthermore, Table 37 illustrates the hyperparameters recommended for model refinement by Bayesian Optimization. Finally, Table 38 presents the aggregated prediction error metrics for both training and testing datasets, calculated over 20 iterations, for the exchanges Kraken, Coinbase, and Binance. The results offer a comprehensive assessment of the model's robustness and predictive capability.

Table 36. List of Selected Features from the Three-Step Feature Selection Process.

Selected Feature	Description
OBV	On-Balance Volume, a momentum indicator that uses volume flow to predict price changes.
trend_dpo	Detrended Price Oscillator, removing the trend from price data.
volatility_bbm	Bollinger Bands Middle line, indicating the average price.
Coin_ETH	Indicator for Ethereum cryptocurrency.
tweets	Number of Twitter mentions or tweets related to the cryptocurrency.
MFI	Money Flow Index, an oscillator that measures buying and selling pressure.
Coin_DASH	Indicator for DASH cryptocurrency.
trend_visual_ichimoku_a	Visual representation of the Ichimoku Cloud's 'A' line.
google_trends	Search interest or volume for a cryptocurrency based on Google Trends data.
momentum_kama	Kaufman's Adaptive Moving Average, a moving average considering market noise.
github_commits	Number of commits or updates made to a cryptocurrency's GitHub repository.
WilliamsR	Williams %R, a momentum indicator measuring overbought and oversold levels.
MACD_hist	MACD histogram, indicating the difference between the MACD line and the Signal line.
momentum_stoch_rsi	Stochastic RSI, an oscillator that measures the RSI's level relative to its high-low range.
VWAP	Volume Weighted Average Price, the average price weighted by volume.
daily_active_addresses	Number of unique addresses involved in transactions on a given day.
trend_ichimoku_conv	Ichimoku Conversion Line, average of the highest high and the lowest low.
ADX	Average Directional Index, an oscillator measuring the strength of a prevailing trend.
volume_mfi	Volume-adjusted Money Flow Index.
Taker Fees	Fees charged for accepting an offer to buy a cryptocurrency.
momentum_rsi	Momentum based on the Relative Strength Index.
Coin_BTC	Indicator for Bitcoin cryptocurrency.

AML_KYC_required	Indicates if Anti-Money Laundering (AML) and Know Your Customer (KYC) verifications are required for transactions.
RSI	Relative Strength Index, a momentum oscillator measuring speed and change of price movements.
MACD_signal	Signal line derived from the MACD to trigger buy/sell signals.
bid_ask_spread	Difference between the highest bid price and the lowest ask price in an order book.
Coin_XRP	Indicator for Ripple (XRP) cryptocurrency.
Maker Fees	Fees charged for making an offer to sell a cryptocurrency.
volatility_bbh	Bollinger Bands High line, indicating the upper price threshold.
transaction_volume	Total volume of cryptocurrency transactions over a specific period.

This table presents the features retained after a rigorous selection process described in Algorithm 4. The initial step employs the Boruta method on the input data. Subsequently, highly correlated features are filtered out. Lastly, features demonstrating high multicollinearity are eliminated.

Table 37. Hyperparameters suggested by Bayesian Optimization

Hyperparameter Name	Hyperparameter Value
Learning rate	0.0067
Dropout rate	0.126
Count of hidden layers	3
Neuron count per layer	459
Batch size	32
Activation function	relu
Optimizer	Adamx
CNN filters	128
Kernel size	3
Pooling size and type	2
Gradient Clipping	5.2
Weight Decay (L2 reg.)	0.0023

Source: Self Compilation.

Table 38. Price Inconsistency Prediction Errors Across Major Cryptocurrency Exchanges: Kraken, Coinbase, and Binance.

Dataset	Exchange	RMSE	MAE	MAPE
Train	**Kraken**	1.2734	1.0716	1.4648
Test	**Kraken**	1.0720	1.0032	1.2720
Train	**Coinbase**	3.3227	3.0226	4.1032
Test	**Coinbase**	3.0194	2.9219	3.8320
Train	**Binance**	2.2062	1.9810	2.7064
Test	**Binance**	2.6568	2.2720	2.9556

Source: Self Compilation.

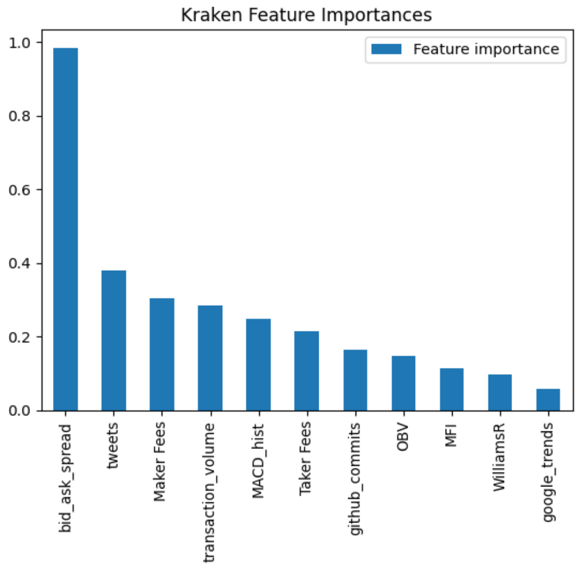
Table 38 illustrates the predictive performance of our model across major exchanges Kraken, Coinbase, and Binance. Low error metrics are observed in both training and testing datasets indicating that the model does not overfit and generalizes reliably to new data. Kraken with the lowest errors demonstrates an efficient market with relatively constrained price inconsistencies. Meanwhile, greater volatility in Coinbase and Binance prices manifests in slightly elevated inconsistency prediction errors. However, retained efficacy on unseen test data verifies model resilience. The similar gap between train and test errors provides further evidence that the model accurately captures inherent exchange dynamics without losing generalizability. These errors validate capturing common drivers of pricing deviations including liquidity, fees, spreads, maturity etc. Impact likely scales with market fragmentation. These results show the importance of exchange-specific characteristics in price inconsistency prediction and the model's capacity to adapt to diverse market environments.

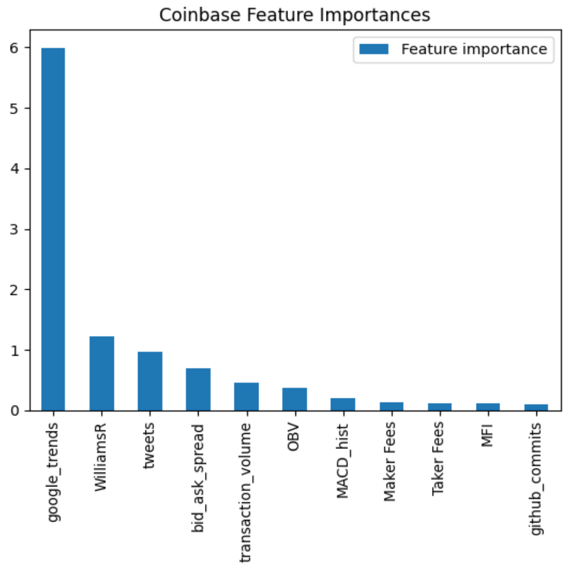
The SHAP feature importance results, as presented in Table 39 and Figures 12 (a), (b), and (c), provide a detailed understanding of the factors contributing to price inconsistencies across the three major cryptocurrency exchanges: Kraken, Coinbase, and Binance.

Table 39. SHAP Feature Importance for Price Inconsistencies Across Major Cryptocurrency Exchanges: Kraken, Coinbase, and Binance.

Feature Name	Feature Importance (Kraken)	Feature Name	Feature Importance (Coinbase Pro)	Feature Name	Feature Importance (Binance)
bid_ask_spread	0.984	google_trends	5.9928	tweets	0.8822
tweets	0.3803	WilliamsR	1.2191	bid_ask_spread	0.7355
Maker Fees	0.3029	tweets	0.9598	transaction_volume	0.6326
transaction_volume	0.2833	bid_ask_spread	0.6869	Maker Fees	0.4463
MACD_hist	0.248	transaction_volume	0.4525	MACD_hist	0.3852
Taker Fees	0.2148	OBV	0.364	Taker Fees	0.1939
github_commits	0.165	MACD_hist	0.1911	RSI	0.1292
OBV	0.1485	Maker Fees	0.1261	github_commits	0.1228
MFI	0.1129	Taker Fees	0.1227	ADX	0.0935
WilliamsR	0.0958	MFI	0.1099	google_trends	0.083
google_trends	0.0582	github_commits	0.092	OBV	0.0554
RSI	0.0492	ADX	0.075	WilliamsR	0.0501
ADX	0.0246	RSI	0.0396	MFI	0.0444

Source: Self Compilation.





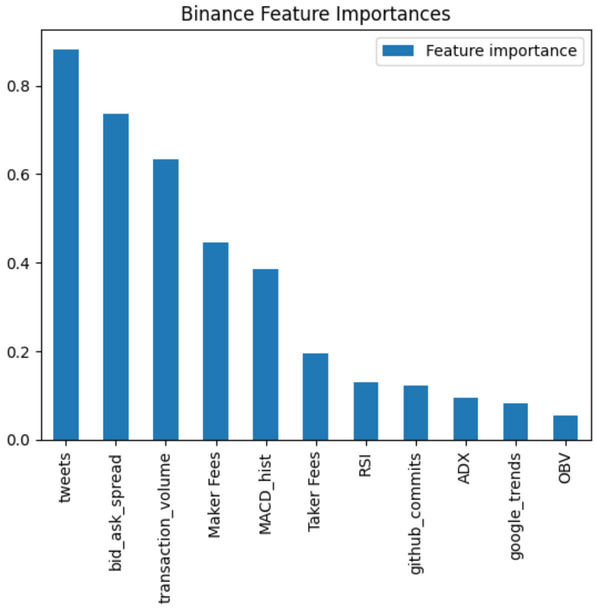


Figure 28: SHAP Feature Importance Plots for Price Inconsistencies Across Major Cryptocurrency Exchanges: (a) Kraken, (b) Coinbase Pro, and (c) Binance (Global)

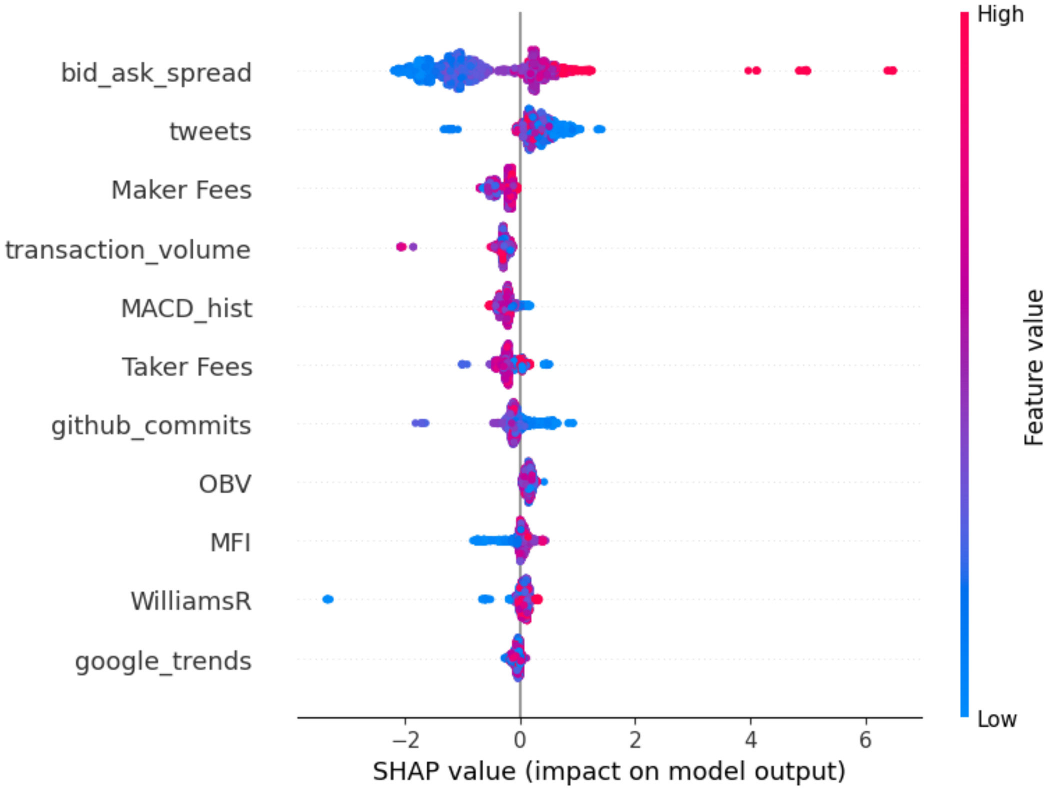


Figure 29 (a). Global Feature Importance for Kraken

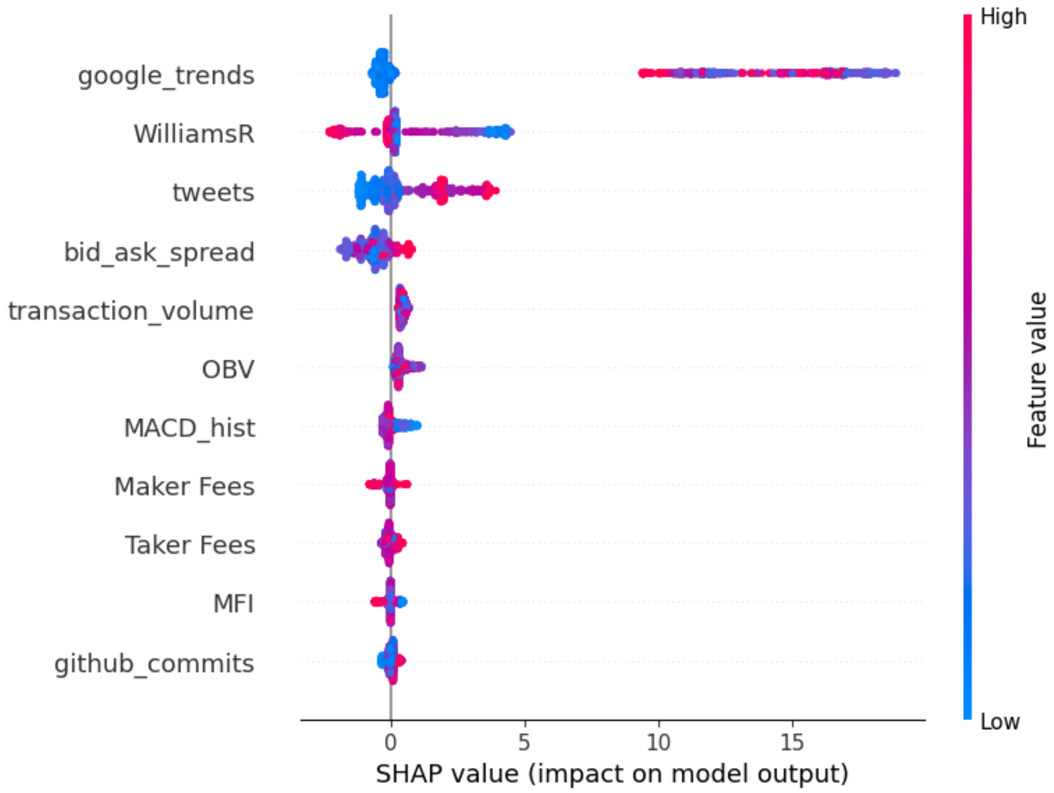


Figure 29 (b). Global Feature Importance for Coinbase Pro

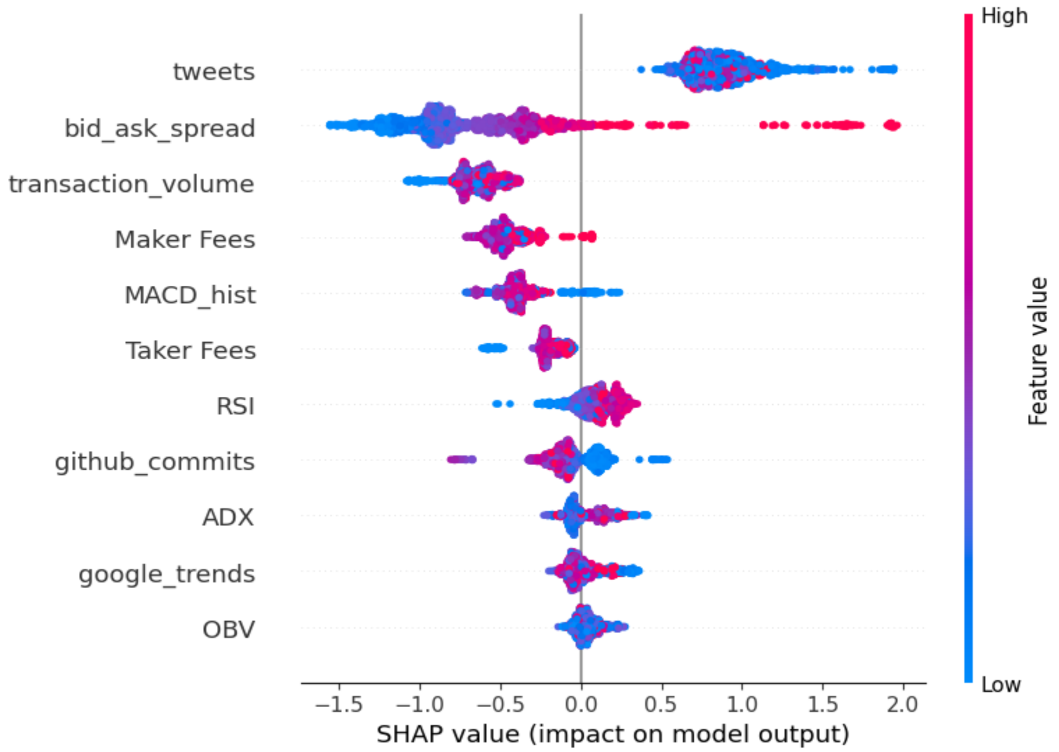


Figure 29 (c). Global Feature Importance for Binance (Global). The graph presents individual data points from the test dataset. On the right, the 'feature value' color scale differentiates values: red signifies higher values, and blue represents lower values. A predominance of red points near the central line indicates a positive correlation with the predicted output, while blue points suggest a negative correlation.

4.3.3.1 Common Factors Influencing Price Inconsistencies Across Kraken, Binance, and Coinbase Pro

As seen in Table 39, Figures 28 (a), (b), and (c), and Figures 29 (a), (b), and (c), the SHAP values for tweets are consistently high across Kraken (0.3803), Coinbase Pro (0.9598), and Binance (0.8822). The influence of tweets or social media sentiment is prominently visible for all three exchanges. The decentralized nature of cryptocurrencies makes them highly susceptible to public sentiment. Any news, regulatory updates, or influential figures' opinions can trigger massive buy or sell actions, leading to price discrepancies, especially if exchanges react at different speeds. Moreover, the interconnectedness and real-time nature of social media platform with users, generates a cascading effect, amplifying the impact of sentiment on price fluctuations.

Upon examining the SHAP feature importance values for Kraken, Binance, and Coinbase Pro, several findings emerge:

1. **Bid-Ask Spread**: The Bid-Ask spread is consistent across all exchanges, indicating its fundamental role in price determination and consistency. The Bid-Ask spread SHAP values for Kraken (0.984), Coinbase Pro (0.6869), and Binance (0.7355). The Bid-Ask spread serves as an immediate reflection of market liquidity and the equilibrium between demand and supply on an exchange. Differences in spread values across exchanges can emanate from variations in liquidity and the nature of traders on each platform. An exchange with a more diverse user base or higher institutional participation may exhibit different liquidity patterns, leading to variances in the bid-ask spread. Such disparities can subsequently signal potential arbitrage opportunities for traders who can leverage price differentials between exchanges.
2. **Social Media Influence (Tweets)**: The influence of social media, specifically Twitter, is evident across all three major exchanges, as corroborated by the SHAP values: Kraken (0.3803), Binance (0.8822), and Coinbase Pro (0.9598). These figures indicate that discourse and sentiment on Twitter have a direct impact on price determinants for these cryptocurrency platforms.
3. **Transaction Volume**: Transaction volume, as shown in Figures 28 (a), (b), and (c), is a significant determinant of price discrepancies, especially on Binance and Kraken. Changes in transaction volumes can lead to price fluctuations. Exchanges with varied user demographics, particularly those with more institutional or large-scale traders, might experience distinct volume changes in short durations, causing temporary price differences between exchanges.
4. **Maker Fees**: Maker Fees, representing charges for placing a new order, play a role in price determination across the exchanges. The SHAP values indicate their relevance on Kraken (0.3029), Binance (0.4463), and to a lesser extent on Coinbase Pro. The imposed

fee structure can influence traders' decisions and subsequently affect the prevailing prices on these platforms.

5. **Technical Indicators (MACD_hist)**: The MACD histogram, a technical indicator in trading, impacts price inconsistencies on Kraken (0.2480), Binance (0.3852), and Coinbase Pro (0.1911). As traders use this indicator in their strategies, changes in the MACD histogram can influence trading patterns and subsequently, price variations.
6. **Other Financial Metrics**: On-Balance Volume (OBV) is observed to have a role in price inconsistencies on Kraken (0.1485) and Coinbase Pro (0.3640). Similarly, the Relative Strength Index (RSI) impacts Binance with a value of 0.1292. Both OBV and RSI serve as indicators of market momentum and can influence the trading choices of investors, resulting in price fluctuations.

Interestingly, features like the MACD histogram, OBV, and the number of GitHub commits are significant across multiple exchanges, indicating their general importance in cryptocurrency price formation. Conversely, attributes like daily active addresses, MACD signal, AML/KYC mandates, and distinct coin indicators such as Coin_BTC, Coin_DASH, Coin_ETH, and Coin_XRP exhibit no significance across all platforms, implying their limited influence on the observed price inconsistencies.

4.3.3.2 Factors Influencing Price Inconsistencies on Kraken

The Bid-Ask spread holds the highest importance in determining price inconsistencies on Kraken, followed by the number of tweets and the maker fees, as shown in Figure 28(a). This suggests that the price inconsistencies on Kraken are primarily driven by the difference between the highest price that a buyer is willing to pay and the lowest price that a seller is willing to accept (bid-ask spread), social media sentiment (tweets), and the fees charged to the maker in a transaction.

1. **Bid-Ask spread**: The Bid-Ask spread is important in influencing price discrepancies on Kraken, as reflected by its SHAP value of 0.984. A wider bid-ask spread frequently indicates limited liquidity or heightened market uncertainty. These characteristics on Kraken, when compared to other exchanges, suggest potential differences in market conditions and participant behavior, which contribute to price disparities.
2. **Tweets**: The SHAP value for tweets on Kraken is 0.3803, indicating the influence of social media sentiment on its cryptocurrency prices. It suggests that Kraken's market participants may be more sensitive to information or events circulated on platforms like Twitter, resulting in price adjustments that can differ from other exchanges.
3. **Maker Fees**: The SHAP value for Maker Fees on Kraken is 0.3029, highlighting their influence on price discrepancies. Elevated maker fees deter market-making activities,

which in turn affects the efficiency in price discovery on Kraken. This directly results in observed price disparities when compared to other exchanges.

4.3.3.3 Factors Influencing Price Inconsistencies on Coinbase Pro

On Coinbase, Google trends data is the most significant feature, indicating that public interest in cryptocurrencies, measured by Google search volume, plays a crucial role in price formation on this exchange. This is followed by the Williams %R (a technical trading indicator that shows whether a security is overbought or oversold), and the number of tweets, as depicted in Figure 28 (b).

1. The SHAP value for Google Trends on Coinbase Pro is 5.9928, highlighting its role in determining price discrepancies. Google search trends can reflect prevailing public sentiment or interest regarding cryptocurrencies.
2. According to Figure 28(b), the Williams %R registers a significant feature importance with a SHAP score of 1.2191. The elevated SHAP score of Williams %R highlights its role in driving price disparities on the platform. Its quick response to market changes can lead to short-lived pricing differences between exchanges, especially when market sentiments to specific events diverge. A consistently oversold reading from Williams %R might indicate reduced buying interest and limited liquidity on Coinbase Pro. These factors can lead to distinct price disparities compared to other exchanges with stable trading activities.
3. Additional contributors to price variations on Coinbase Pro include tweets, the Bid-Ask Spread, and Transaction Volume.

4.3.3.4 Factors Influencing Price Inconsistencies on Binance

For Binance, the number of tweets is the most important feature, followed by the bid-ask spread and the transaction volume, as illustrated in Figure 28(c). This suggests that social media sentiment, market liquidity, and the difference between the highest bid and the lowest ask prices are key drivers of price inconsistencies on Binance.

1. On Binance, tweets related to specific cryptocurrencies influence price inconsistencies, evidenced by a SHAP value of 0.8822. Binance has an extensive following on Twitter, with over 10 million followers, more than Coinbase Pro and Kraken. This large following implies that information on Twitter can have a considerable impact on Binance's market, contributing to the price variations.
2. On Binance, the Bid-Ask spread is important in price disparities, evidenced by a SHAP value of 0.7355. Binance, one of the world's premier cryptocurrency exchanges in terms of volume, can encounter liquidity constraints during periods of heightened volatility, which is frequent in cryptocurrency markets. This can result in broader bid-ask spreads, contributing to the observed price discrepancies.

3. Transaction volume is another influential factor in determining price inconsistencies on Binance, with a SHAP score of 0.6326. Elevated transaction volumes can infuse volatility into the market, particularly when significant trading activities are condensed into brief periods. Such conditions can result in momentary mispricings, contributing to the inconsistencies witnessed on the platform.

The SHAP summary plots, as represented in Figures 28 (a), (b), and (c) for Kraken, Coinbase Pro, and Binance respectively, visually corroborate the findings derived from the SHAP feature importance values. These plots provide a comprehensive visualization of the impact and directionality of each feature on the model's output. For Kraken, as shown in Figure 28 (a), the bid-ask spread emerges as a dominant factor, reinforcing its significance in determining price inconsistencies. Similarly, in Figure 28 (b) for Coinbase Pro, the predominance of Google Trends data is evident, reflecting the exchange's sensitivity to public sentiment and interest. Figure 28 (c) for Binance showcases tweets as an important influence, emphasizing the role of social media sentiment in the platform's price formation.

The results of our analysis, as presented in Table 39, Figures 28 (a), (b), and (c), and Figures 29 (a), (b), and (c) provide a detailed understanding of the factors contributing to price inconsistencies across the three major cryptocurrency exchanges: Kraken, Coinbase, and Binance.

4.3.4 Implications for Stakeholders: Traders, and Arbitrageurs

The SHAP feature importance analysis provides a detailed understanding of the factors influencing price inconsistencies across cryptocurrency exchanges. The significant impact of social media sentiment, especially from tweets and Google trends data across all exchanges, indicates the role of public sentiment and interest in shaping cryptocurrency prices. This suggests that investors and traders might use such publicly available data to predict potential price movements.

Market liquidity, as reflected by the importance of bid-ask spread and transaction volume, plays a central role in price discrepancies across exchanges. This observation is particularly relevant for high-frequency traders and those seeking to benefit from price differentials. Features like daily active addresses, MACD signal, AML/KYC requirements, and specific coin indicators have minimal importance across all exchanges. Anti-Money Laundering (AML) and Know Your Customer (KYC) protocols underwent significant evolution post the 2017 cryptocurrency market crash. Governments worldwide began instituting stricter regulatory frameworks, necessitating almost all users on platforms like Kraken, Binance, and Coinbase Pro to undergo KYC procedures before transacting. This led to a more transparent and structured market environment. However, by the time our dataset was compiled, these regulations had largely been integrated into the standard operating procedures of these exchanges. The user base had adjusted to these requirements. As a result, while AML and KYC played a pivotal role in shaping the crypto landscape and could have influenced user onboarding and overall market trust, their day-to-day influence on price inconsistencies might have diminished. This is because the markets had largely absorbed and adjusted to these regulatory norms, rendering them as constants rather than variables that would contribute to short-term price volatility. Hence, while AML and KYC are fundamental to the operation and credibility of exchanges, their direct impact on intra-day or short-term price inconsistencies across Kraken, Binance, and Coinbase Pro seems to be minimal in the present context.

These observations can help researchers and market analysts focus on more relevant factors when analyzing cryptocurrency price behavior. The distinct importance of features unique to each exchange, such as Maker and Taker fees, highlights the different characteristics and operations of each platform. Recognizing these differences is essential for developing effective trading strategies and research methods.

4.4 Objective 4: Liquidity Comparison

0bjective 4: *To compare the liquidity of cryptocurrency exchanges with the liquidities of different sizes of selected stocks by using Martin Liquidity Index (MLI)*

4.4.1 Liquidity Analysis: Metrics and Comparisons Across Exchanges

In this section, we present the results of computed liquidities using Algorithm 5 from Chapter 3. We evaluate several liquidity measures, with the Martin Liquidity Index (MLI) as our primary metric. Additionally, we consider measures such as Amihud's Illiquidity Ratio (AIR), AR Bid-Ask Spread, Roll's Covariance Liquidity Estimator, and the CS spread estimator. Our analysis encompasses three major cryptocurrency exchanges: Binance, Coinbase, and Kraken. We also compare their liquidity measures with traditional stock markets using NYSE, NASDAQ, NIFTY, and BSE SENSEX data. Tables 40, 41, and 42 present the computed liquidities for each market across the selected measures for Intervals 1, 2, and 3, respectively. A key feature in these tables is the rank column for each liquidity measure, which indicates the liquidity position of each exchange. A rank of 1 denotes the highest liquidity, with higher numbers indicating reduced liquidity. Adding a rank column offers clarity on the liquidity position of each market. Additionally, the multiple intervals offer a temporal view, highlighting how liquidity levels evolve in each market. Figures 9(a) to 9(e) provide visual representations of liquidity measures across exchanges using radar charts, aiding in the comparative analysis of market liquidity for the discussed intervals.

Table 40. Liquidity Measures Across Major Cryptocurrency and Traditional Markets for Interval 1.

The rank for each measure indicates the liquidity position of the exchange relative to others, with a rank of 1 being the most liquid and higher ranks indicating lower liquidity.

Exchange	MLI Value	MLI Rank	Amihud	Amihud Rank	AR Bid-Ask Spread	AR Rank	CS Estimator	CS Rank	Rolls Estimator	Rolls Rank
NASDAQ100	1.2874E-04	1	2.8026E-12	2	2.8056E+02	1	8.2423E-05	2	2.9076E-10	2
NYSE Composite	5.2687E-04	2	1.4219E-12	1	3.7922E+02	2	6.0441E-05	1	2.4964E-10	1
NIFTY	2.6878E-02	3	3.4945E-10	3	3.8614E+02	3	1.3147E-04	3	3.4528E-10	3
BSE SENSEX	6.9096E+02	4	3.6524E-07	4	1.6988E+03	4	1.4214E-04	4	3.6256E-10	4
Binance	6.9187E+02	5	4.0812E-07	5	1.7516E+03	5	1.6667E-04	5	3.7985E-10	5
Kraken	8.9826E+02	6	1.0850E-06	6	2.1777E+03	6	1.9168E-04	6	6.9713E-10	6
Coinbase Pro	1.1056E+03	7	1.4599E-06	7	2.6039E+03	7	2.1670E-04	7	7.1441E-10	7

Source: Author’s Calculations**.**

Table 41. Liquidity Measures Across Major Cryptocurrency and Traditional Markets for Interval 2

Exchange	MLI Value	MLI Rank	Amihud	Amihud Rank	AR-Bid Ask Spread	AR Rank	CS Estimator	CS Rank	Rolls Estimator	Rolls Rank
NYSE Composite	5.7197E-03	1	1.9111E-12	1	6.3122E+02	1	2.4590E-04	2	1.1000E-11	1
NASDAQ100	7.6459E-03	2	2.9147E-12	2	7.8898E+02	2	2.4128E-04	1	1.4470E-11	2
NIFTY	3.7788E+01	3	1.0906E+05	3	9.2599E+02	3	4.1167E-04	4	2.7840E-10	3
Binance	1.6354E+02	4	2.2246E+07	4	2.6659E+03	4	2.7039E-04	3	3.9010E-10	4
BSE SENSEX	1.8243E+02	5	2.2661E+07	5	2.8133E+03	5	4.2392E-04	5	5.2380E-10	5
Coinbase Pro	2.3527E+02	6	5.1906E+07	6	3.4374E+03	6	4.7657E-04	6	8.2743E-10	6
Kraken	2.8811E+02	7	6.3029E+07	7	4.0615E+03	7	5.2921E-04	7	1.1701E-09	7

Source: Author’s Calculations**.**

Table 42. Liquidity Measures Across Major Cryptocurrency and Traditional Markets for Interval 3

Exchange	MLI Value	MLI Rank	Amihud	Amihud Rank	AR Bid-Ask Spread	AR Rank	CS Estimator	CS Rank	Rolls Estimator	Roll's Rank
NYSE Composite	2.0410E-01	1	1.9889E-12	1	6.0988E+02	1	2.3166E-04	1	3.5671E-10	1
NASDAQ100	2.6753E+00	2	2.9109E-12	2	8.3418E+02	2	2.5928E-04	2	4.9183E-10	2
NIFTY	3.0788E+01	3	9.9774E+04	3	8.6840E+02	3	2.6588E-04	3	8.6883E-10	3
BSE SENSEX	7.6780E+01	4	2.0352E+07	4	2.5829E+03	4	3.8223E-04	4	1.0323E-09	5
Binance	2.9358E+02	5	2.5503E+07	5	2.8012E+03	5	3.9934E-04	5	8.9717E-10	4
Kraken	3.9513E+02	6	4.2936E+08	6	3.1615E+03	6	4.9100E-04	7	1.1203E-09	6
Coinbase Pro	6.7580E+02	7	6.6274E+08	7	3.8058E+03	7	4.4517E-04	6	1.1674E-09	7

Source: Author’s Calculations**.**

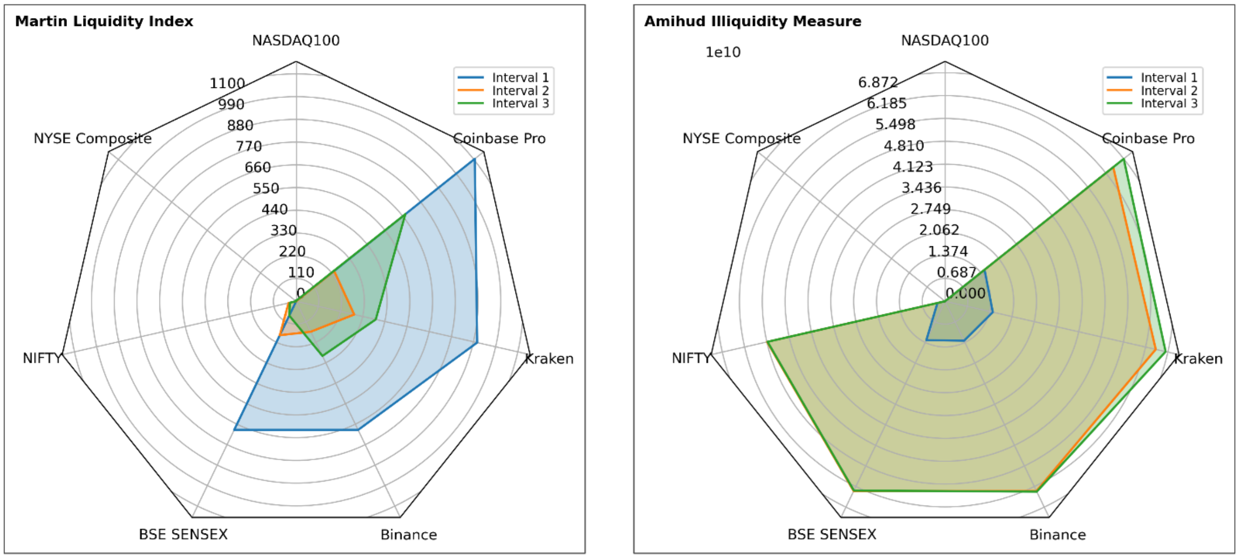


Figure 30 (a) and 30 (b). Radar chart representation of the Martin Liquidity Index (MLI) and Amihud Illiquidity Measure values across various exchanges for all three intervals. Figure 14 (a) showcases the MLI, while Figure 14 (b) illustrates the Amihud Illiquidity Measure, transformed using a logarithmic scale for clarity.

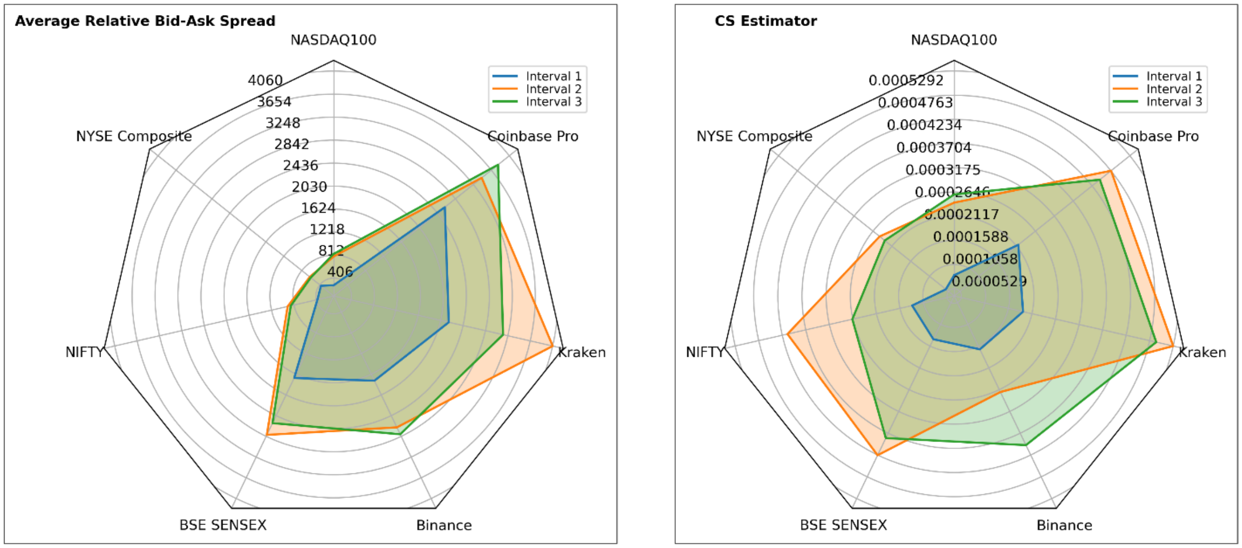


Figure 30 (c) and 30 (d). Radar chart representation of the Average Relative Bid-Ask Spread and CS Estimator values across different exchanges for three distinct intervals. Figure 30 (c) depicts the Average Relative Bid-Ask Spread, and Figure 30 (d) presents the CS Estimator.

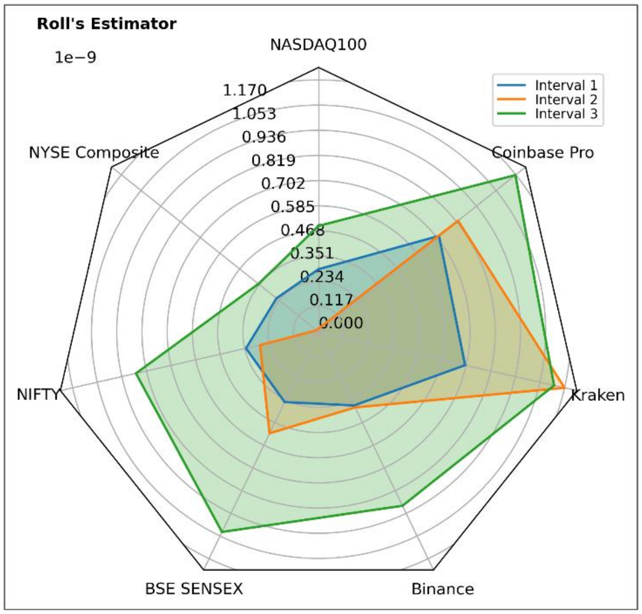


Figure 30 (e). Radar chart representation of Roll's Estimator for the exchanges, presented with a scaled axis to accommodate the range of values, marked at 1×10^{-9}.

4.4.1.1 Liquidity Analysis for Interval 1

According to Table 40, traditional exchanges such as NASDAQ100, NYSE Composite, and NIFTY consistently display higher liquidity metrics than BSE SENSEX and the cryptocurrency exchanges. Specifically, NASDAQ100 holds the most favorable MLI value of 1.2874×10^{-4}, indicating it as the most liquid exchange based on our baseline measure. It is closely followed by NYSE Composite and NIFTY, with MLI Ranks of second and third, respectively. Although liquid with an MLI value of 6.9096×10^{2}, BSE SENSEX ranks fourth, lagging the major traditional stock markets. For the Amihud Illiquidity Measure, NYSE Composite, securing the top position with a first rank, emerges as the most liquid, closely trailed by NASDAQ100, which holds the second position. BSE SENSEX's performance in this measure places it in the fourth position, indicating lower liquidity than NASDAQ100 and NYSE Composite but higher than the cryptocurrency exchanges. In the AR Bid-Ask Spread measure, NASDAQ100, at the premier position with a first rank, showcases minimal spread and higher liquidity. NYSE Composite and NIFTY follow suit with second and third ranks, respectively, while BSE SENSEX ranks fourth. NYSE Composite ranks first based on the CS Estimator and Rolls Estimator, indicating its leading liquidity. NASDAQ100 and NIFTY follow with second and third positions, respectively. BSE SENSEX consistently ranks fourth in both measures. These rankings suggest that NASDAQ100 and NYSE Composite are the most liquid markets, with minimal bid-ask spreads. Despite being slightly lower than the top two exchanges, NIFTY also demonstrates good liquidity. On the other hand, BSE SENSEX consistently lags in terms of liquidity, as it ranks fourth in both measures.

Binance demonstrates the highest liquidity among the cryptocurrency exchanges, closely followed by Kraken and then Coinbase Pro, as evident from their respective ranks across all five liquidity measures. However, the liquidity of Binance remains below the levels of traditional exchanges. Liquidity indicators, such as the Martin Liquidity Index and the Rolls Estimator, substantiate these observations. For instance, in the Martin Liquidity Index (MLI), Binance's value stands at 6.9187×10^{2}, followed by Kraken with 8.9826×10^{2} and Coinbase Pro at 1.1056×10^{3}. The Amihud Illiquidity Measure, for instance, shows the significant discrepancies between regular trading platforms and cryptocurrency exchanges, emphasizing the more established nature of traditional exchanges. This sequence is mirrored across measures like the Amihud Illiquidity Measure, where Binance, Kraken, and Coinbase Pro have values of 4.0812×10^{-7}, 1.0850×10^{-6}, and 1.4599×10^{-6}, respectively. While Binance exhibits the highest liquidity among its cryptocurrency counterparts, Kraken and Coinbase Pro demonstrate an observable liquidity discrepancy. The differences in their values across the measures, such as the MLI and Amihud Illiquidity Measure, suggest Kraken has marginally better liquidity than Coinbase Pro, although both trail Binance. During Interval 1, the cryptocurrency market was recovering from the 2018 crash (Lahmiri & Bekiros, 2019b). Binance, initially founded in China, relocated its operations to countries like Japan and Malta in response to intensifying regulatory pressures from the Chinese government (Disli et al., 2022). This strategic move allowed Binance to sustain its operations in environments with more accommodating regulations. Meanwhile, traditional stock exchanges such as NASDAQ100 and NYSE Composite consistently demonstrated high liquidity, largely due to their established nature and reduced susceptibility to crypto-specific events.

Table 40 further presents a detailed comparison between Binance and BSE SENSEX, highlighting the evolving financial landscape. Specifically, the MLI values for BSE SENSEX and Binance are nearly identical at 6.9096×102 and 6.9187×102, respectively, securing them the fourth and fifth positions. This proximity continues to the Amihud Illiquidity Measure, with BSE SENSEX at 3.6524×10−7 and Binance following closely at 4.0812×10−7, again placing them fourth and fifth. The CS Estimator and Rolls Estimator metrics reinforce this pattern, emphasizing the comparable liquidity profiles of the two markets during Interval 1.

While the liquidity metrics of Binance and BSE SENSEX appear closely aligned in this instance, it is important to consider the context of this observation. During this specific interval, Binance's liquidity metrics closely resemble BSE SENSEX's. However, Binance exhibits significantly lower liquidity than other established traditional exchanges. For Interval 1, Binance implemented several strategic initiatives to enhance its market position. The cryptocurrency exchange launched its futures trading platform, allowing users to trade futures contracts on various cryptocurrencies with up to 125x leverage (Kawai et al., 2023). Simultaneously, the platform introduced a lending mechanism, enabling cryptocurrency holders to earn interest, and expanded its fiat gateways to cater to a global audience. Additionally, initiating a staking platform and including margin trading provided further avenues for user engagement and potential liquidity enhancement. In stark contrast, the BSE SENSEX, which represents 30 established Indian companies, typically offers a trading environment that is more stable and predictable than cryptocurrencies (Soni, 2012). Such observations underscore the evolving financial landscape, where, under certain conditions, digital assets can exhibit liquidity metrics on par with traditional markets. Under the given conditions of Interval 1, these findings suggest that digital assets can exhibit liquidity metrics that align closely with those of traditional markets.

4.4.1.2 Liquidity Analysis for Interval 2

Table 41 presents the liquidity comparison results across cryptocurrency and traditional markets for Interval 2. During this period, the NYSE Composite and NASDAQ100 continued to dominate the liquidity rankings, securing the top two positions across almost all measures. The NYSE Composite, with an MLI value of 5.7197×10^{-3} and an Amihud value of 1.9111×10^{-12}, stands as the most liquid exchange during this period. NASDAQ100 follows closely, registering an MLI value of 7.6459×10^{-3} and an Amihud value of 2.9147×10^{-12}. Both exchanges maintain the highest liquidity in the AR Bid-Ask Spread measure and the Rolls Estimator. Interestingly, while the NYSE Composite leads in most measures, NASDAQ100 marginally surpasses it in the CS Estimator. NIFTY maintained its third position across all measures, like the previous interval. Its liquidity metrics, such as an MLI value of 3.7788×10^{1} and an Amihud value of 1.0906×10^{5}, reflected its consistent performance relative to other exchanges. Binance continued to lead among cryptocurrency exchanges, ranking fourth in most measures. With an MLI value of 1.6354×10^{2} and an Amihud value of 2.2246×10^{7}, Binance's liquidity profile remained ahead of other cryptocurrency exchanges but mostly below the major traditional exchanges.

The liquidities of Coinbase Pro and Kraken were ranked sixth and seventh positions, respectively. Their values, especially in the MLI and Amihud measures, indicate a discernible gap in liquidity compared to both traditional exchanges and Binance. Specifically, with its MLI value of 2.8811×10^{2} and Amihud value of 6.3029×10^{7}, Kraken ranked last, suggesting it had the least liquidity among the studied exchanges during Interval 2.

Notably, Binance surpassed BSE SENSEX in this interval, which stood fifth in the rankings. The SP BSE SENSEX, with its MLI value of 1.8243×10^{2} and an Amihud value of 2.2661×10^{7}, exhibited liquidity metrics closely aligned with Binance but slightly lower. Binance displayed significant improvements in liquidity metrics during this interval. These improvements could be attributed to the significant growth in the cryptocurrency market, as evidenced by the total market capitalization of cryptocurrencies escalating from roughly $190 billion at the start of this period to a staggering $2.9 trillion by November 2021. Such rapid growth, exemplified by Bitcoin's peak value of approximately $68,789 in November 2021, invariably draws trading volumes, thereby enhancing liquidity (Aljadani, 2022). Institutional investments in the crypto space further bolstered this trend. Firms like MicroStrategy, Tesla, and Square made notable investments in Bitcoin, legitimizing cryptocurrencies for many conventional investors and injecting substantial trading volumes (Field & Inci, 2023). This influx undoubtedly bolstered the liquidity of platforms like Binance. As a result, Binance experienced increased trading activity and attracted a wider range of traders, including retail and institutional investors. The enhanced liquidity on Binance provided traders better opportunities for executing trades at desired prices, ultimately contributing to the platform's overall success in the crypto market.

On the other hand, BSE Sensex underwent a turbulent phase during the COVID-19 pandemic, primarily due to supply chain disruptions. Such disruptions can lead to a decline in stock valuation and, consequently, liquidity (S. Kumar et al., 2015). During this period, major Indian companies faced significant supply chain challenges due to lockdowns. The events such as the sharp declines in February 2020 and the market reactions following the World Health Organization's pandemic declaration (https://www.who.int/europe/emergencies/situations/covid-19) reflected the uncertainties shadowing global investors. Coupled with the shrinkage in the total market cap by over 27.31% from the beginning of the year, these factors influenced its liquidity. Overall, Interval 2 reveals a shifting liquidity landscape with traditional exchanges maintaining their dominance, Binance edging out BSE SENSEX, and a noticeable liquidity discrepancy between the top cryptocurrency exchange and its counterparts.

4.4.1.3 Liquidity Analysis for Interval 3

Table 42 combines data from Intervals 1 and 2, from 08-Sep-19 to 01-Apr-23. This period provides a comprehensive view of liquidity trends across both traditional and cryptocurrency exchanges. NYSE Composite and NASDAQ100 consistently occupy the top spots, indicating their strong liquidity. The NYSE Composite has an MLI value of 2.0410×10^{-1} and an Amihud measure of 1.9889×10^{-12}. The NASDAQ100 follows with an MLI of 2.6753 and an Amihud of 2.9109×10^{-12}. These exchanges also lead to the AR Bid-Ask Spread measure and the Rolls Estimator. NIFTY maintains its third position with an MLI of 3.0788×10^{1} and an Amihud of 9.9774×10^{4}. This ranking is consistent with previous intervals. SP BSE SENSEX ranks fourth, showing improved liquidity compared to Interval 2. This change is likely due to the Indian economy's recovery after the pandemic (Jha & Jha, 2020). The MLI for BSE SENSEX is 7.6780×10^{1}, and its Amihud value is 2.0352×10^{7}. Binance is fifth, with an MLI of 2.9358×10^{2} and an Amihud of 2.5503×10^{7}. While it leads among cryptocurrency exchanges, it lags major traditional exchanges. Kraken and Coinbase Pro are in the sixth and seventh positions. Their rankings suggest a liquidity gap compared to Binance and traditional exchanges.

Complementing the tabular findings, Figures 30 (a) to 30 (e) visually illustrate exchange liquidities using radar charts across exchanges and intervals. These charts restate the consistent top-tier liquidity of traditional exchanges like NYSE Composite and NASDAQ100 across all intervals. Binance's rise in liquidity, especially in Interval 2, and its subsequent position relative to BSE SENSEX is also evident. The main findings from the analysis across Intervals 1, 2, and 3 indicate that traditional exchanges, notably NASDAQ100, NYSE Composite, and NIFTY, consistently maintain robust liquidity metrics throughout the studied durations. Binance stands out within cryptocurrency exchanges, displaying notable liquidity that, at certain junctures, even rivals BSE SENSEX. However, it is crucial to highlight that BSE SENSEX's liquidity demonstrates oscillations, with a significant decline in Interval 2 that could be attributed to global factors such as the COVID-19 pandemic. By Interval 3, a discernible recovery in its liquidity is observed, mirroring the resilience and recovery trajectory of the Indian economy.

4.4.2 Implications for Traders, Researchers, and Arbitrageurs.

The results of this objective highlight the growing significance of cryptocurrency exchanges in global financial markets. With platforms like Binance showcasing liquidity metrics that can, at times, rival traditional stock exchanges, researchers are presented with a fascinating evolution in market dynamics. Observing the rise of exchanges like Binance emphasizes the need for further exploration into what drives liquidity in cryptocurrency markets and how these platforms are reshaping the financial landscape.

The simultaneous evaluation of cryptocurrency and traditional exchanges offers a unique perspective, emphasizing the potential of platforms like Binance to challenge long-standing financial institutions. This can inspire deeper research into the forces shaping this emerging competition and the broader implications for rapidly changing landscape of global trading platforms. The findings equip traders with a clear picture of liquidity across various exchanges. Choosing platforms with consistently higher liquidity ensures smoother and more efficient transaction execution. Liquidity variations, particularly on cryptocurrency exchanges and during significant market events, have implications for trading risk. Traders need to be cognizant of these shifts and adapt their strategies to mitigate potential pitfalls.

The differences in liquidity across cryptocurrency and traditional exchanges open opportunities for arbitrageurs. Recognizing these disparities allows for buying on lower liquidity exchanges and selling on higher ones. Arbitrageurs can benefit from understanding which exchanges consistently offer high liquidity. For instance, focusing on exchanges like Binance during its high liquidity phases can yield more opportunities. The changing liquidity landscape suggests that arbitrageurs need adaptive strategies. Being able to adjust based on current liquidity conditions can be beneficial.

Chapter 5

Conclusions and Future Work

This chapter offers a comprehensive summary and conclusions drawn from this research. In each sub-section, a recapitulation of the specific objective is presented, accompanied by a succinct description of the methodologies employed and the principal results derived. For research questions associated with these objectives, we provide answers and evaluate the corresponding hypotheses. The chapter also outlines the main contributions and novelties of this work. Finally, the chapter discusses the broader implications of this research, highlights its limitations, and proposes potential directions for future investigations in this domain.

5.1 Conclusion and Summary of Main Findings: Objective 1

Objective 1: *To analyze the trends of exchange prices of Bitcoin and Altcoins and to forecast future exchange prices*

5.1.1 Recap of Objective 1: Our objective was to analyze the trends of exchange prices of Bitcoin and Altcoins and to forecast future exchange prices of Bitcoin and Altcoins. We divided our objectives into two parts. We analyzed all cryptocurrencies' trends and statistical properties in the first part. We proposed a novel forecasting framework for future exchange prices in the second part.

5.1.1.1 Overview of Cryptocurrency Trends

In Part One, we first presented the results for analyzing the trends of exchange prices by Performing Change Point Analysis using the PELT algorithm for each of the five-time series. We discussed the statistical properties of the price dynamics for each cryptocurrency. We found the presence of Non-Linear Price Trends and change points of leptokurtosis, which signified the probability of Extreme Price Changes and Asymmetry in price distribution for all five datasets. We also found Significant Fluctuations in Volatility at multiple change points. Next, we conducted a Multifractal Detrended Fluctuation Analysis of all cryptocurrencies under study and discussed the multifractal nature of all series at daily and weekly scales. We found the presence of persistence and multifractality in all five datasets, which highlighted the necessity of creating a sophisticated Forecasting framework that successfully models the non-linear and non-stationary nature of cryptocurrencies.

5.1.1.2 Novel Forecasting Framework

In Part Two, we introduced a novel forecasting framework that employed various combinations of input datasets: fundamental indicators, technical indicators, and their combination. This framework aimed to predict the closing prices of Bitcoin, Ethereum, and other Altcoins over prediction horizons ranging from the next day to 7 days. While our preliminary experiments explored several machine learning models, we ultimately selected three Deep Learning models: CNN-BiLSTM with self-attention, BiLSTM with self-attention mechanism, and ANN. Each model was trained using three distinct input datasets: Fundamental Indicators (FI), Technical Indicators (TI), and a combination of both (FI+TI).

We proposed a novel Signal Processing mechanism to improve the performance of our machine learning models. Outliers were identified and addressed using the Isolation Forest Algorithm. Subsequently, we tackled the challenge of inherent signal noise in the multivariate time series datasets. We employed Signal Smoothing techniques to mitigate this, specifically leveraging the Savitzky-Golay Filters complemented with Fitting Weights. After attenuating the signal noise in all the time series, we performed a three-step feature selection to improve the quality of the input datasets further. In the first step, we performed Feature Selection with the Boruta Method and got the important features based on their Z-scores. Next, we applied Pearson Correlation on the features obtained from the Boruta method to eliminate highly correlated features. Finally, we Filtered out features with High VIF to eliminate residual Multicollinearity.

After preparing the data for training, we tested the three deep-learning models. We applied Bayesian Optimization for fine-tuning, enabling the models to identify patterns across the three types of input data. This process was repeated for each period, with each corresponding to a distinct phase in the cryptocurrency market. In our study, we presented results for nine distinct input data-model combinations. Each experiment was conducted 20 times, and we reported the average results for the next day's closing price predictions for Bitcoin, Ethereum, Ripple, Litecoin, and DASH coins. These predictions spanned four intervals, each representing a unique market phase. To further assess the reliability of our models, we extended our forecasting to 3, 5, and 7-day intervals.

5.1.2 Major Findings and Research Outcomes

5.1.2.1 Research Question 1.1

What are the key statistical properties characterizing the price trends of Bitcoin and Altcoins?

5.1.2.1.1 Summary of Main Findings:

1. **Non-Linear Price Trends:** All cryptocurrencies including Bitcoin, Ethereum, DASH, LTC, and XRP, show significant fluctuations in their mean prices over the observed periods, indicating non-linear price trends. There are distinct periods of price growth and decline, as evidenced by the shifts in mean prices across different segments. For instance, Bitcoin experienced significant price growth transitioning from Change Point 7 to Change Point 8 and a decline from Change Point 9 to Change Point 10. The kurtosis values for Bitcoin range from -1.5188 to 2.9473. A high positive kurtosis at Change Point 15 suggests a leptokurtic distribution, indicating a higher probability of extreme price changes.
2. **Presence of Extreme Price Changes in different market phases:** High kurtosis values in several change points across all cryptocurrencies suggest a leptokurtic distribution, indicating a higher probability of extreme price changes and increased market risk. High kurtosis values in segments such as Bitcoin's Change Point 14 highlight increased market risk.
3. **Asymmetry in Price Changes:** The skewness values across all cryptocurrencies indicate periods of both positive and negative skewness, suggesting that price changes are not symmetrically distributed and there are periods of both bullish and bearish trends.
4. **High Volatility:** It was found that among the five cryptocurrencies, Bitcoin had the highest absolute standard deviation, indicating that it experienced the most significant price fluctuations (volatility). This is consistent with Bitcoin's position as the leading cryptocurrency, which often experiences substantial price movements due to various market factors. When comparing the standard deviations relative to their mean values, it was found that each cryptocurrency has periods of increased volatility.
5. **Multifractal Nature:** The Hurst exponent analysis found the presence of multifractality and complex scaling behavior in all five cryptocurrency prices. The consistent rise of h(q) for negative q values, the dip at q=0, and the plateau for q>1 in Figure 9(f), all highlight the multifractal characteristics of cryptocurrency time series.

5.1.2.1.2 Hypothesis Evaluation:

Table 43. Evaluation of Hypothesis for Research Question 1.1

Hypothesis	Hypothesis Statements	Evaluation	Result
H 1.1a	**H1:** The price trends of Bitcoin and Altcoins are non-linear, with distinct periods of growth and decline characterized by shifts in mean prices.	The findings from Multiple Change point and MFDFA suggest that the price trends of cryptocurrencies are non-linear with distinct periods of growth and decline.	Reject H0, Accept H1
H1.1b	**H1:** The price trends of Bitcoin and Altcoins exhibit leptokurtosis, indicating a higher probability of extreme price changes during certain periods.	High kurtosis values in segments like Bitcoin's Change Point 14 suggest an increased likelihood of extreme price changes.	Reject H0, Accept H1
H1.1c		The data indicates periods of	Reject H0,

H1: The price trends of Bitcoin and Altcoins exhibit skewness, indicating asymmetry in price changes with periods of higher probabilities for both large price increases and decreases.	skewness, suggesting asymmetric price changes in the cryptocurrencies.	Accept H1

Source: Author’s compilation

5.1.2.2 Research Question 1.2

How do these statistical properties change across different phases of the cryptocurrency market's evolution?

5.1.2.2.1 Summary of Main Findings:

1. **Fluctuations in Volatility:**
    i. The standard deviation values (from Table 11 to Table 15) highlight significant fluctuations in price volatility across different market phases.
    ii. For instance, Bitcoin's transition from Change Point 8 to Change Point 9 suggests a period of price stabilization following heightened volatility.
    iii. Ethereum's volatility, as denoted by the standard deviation, spans from 36.6047 to 4189.9529, with the peak volatility observed at 529.4631. Such periods of intense price fluctuations can be attributed to various factors, including market news, technological advancements, and regulatory changes.
    iv. These high fluctuations in volatility underscore both the inherent risks and potential opportunities in the cryptocurrency market.

2. **Persistent Behavior:**
    i. A consistent rise in the generalized Hurst exponent, h(q), as seen in Figure 9(f), for negative q values across daily and weekly scales, indicates significant variations in the time series for all analyzed cryptocurrencies.
    ii. This upward trend signifies a persistent behavior, suggesting a positive correlation between past and future prices.

3. **Long-Range Correlations:**
    i. As observed in Figure 9(f), q values transition from negative to positive, with a pronounced increase in the Hurst exponent h(q), highlighting the larger fluctuations in the series.
    ii. This trend suggests that these fluctuations demonstrate persistent behavior, indicative of long-range correlations or a long-memory process.
    iii. In simpler terms, both minor and significant price shifts in cryptocurrencies are connected over extended periods, reflecting the cryptocurrency market's memory of past events.

4. **Complex Scaling Behavior:**
    1. The Hurst exponent analysis, along with other multifractal measures, reveals the intricate scaling behavior of cryptocurrency prices.
    2. This indicates that cryptocurrency price time series possess multifractal properties, meaning they exhibit complex patterns at various scales. Such multifractality can arise from various internal and external factors influencing the market.

The comparisons with the past literature for Objective 1 are extensively reported in Chapter 4 and the proposed model outperforms the previous state-of-the-art.

5.1.2.2.2 Evaluation of Hypothesis

Table 44. Evaluation of Hypothesis for Research Question 1.2

Hypothesis	Hypothesis Statements	Evaluation	Result
H 1.2a	**H1:** The volatility fluctuates significantly across different segments.	Based on the observed volatility fluctuations, the findings support the H1 hypothesis.	Reject H0, Accept H1
H 1.2b	**H1:** Cryptocurrencies exhibit complex scaling behavior with multifractal properties.	The findings, especially from the Hurst exponent analysis, support the H1 hypothesis.	Reject H0, Accept H1

Source: Author’s compilation

5.1.2.3 Research Question 1.3

How do fundamental indicators, technical indicators, and their combination influence the predictive performance in forecasting the closing prices of Bitcoin, Ethereum, and other Altcoins?

5.1.2.3.1 Summary of Main Findings

1. From our analysis of Tables 16-20, we found that the CNN-BiLSTM model, when fed with a combination of Fundamental Indicators (FI) and Technical Indicators (TI), consistently delivers the most accurate next-day price forecasts for Bitcoin across all intervals.
2. For Altcoins, which include Ethereum, Ripple, Litecoin, and DASH, the CNN-BiLSTM model with the FI+TI combination consistently outperforms other models in Interval 4, which contains the most representative data. However, in Intervals 1 to 3, there are instances where the BiLSTM model with FI+TI showcases superior performance on certain metrics.
3. When forecasting over 3-day, 5-day, and 7-day horizons, the CNN-BiLSTM model with FI+TI consistently exhibits robust performance.

When we compared the performance of our top-performing model with various past studies in the literature, we found that our model consistently outperforms the existing state-of-the-art in most of the performance metrics. These comparisons are discussed in detail in Chapter 4.

5.1.2.3.2 Evaluation of Hypothesis

Table 45. Evaluation of Hypothesis for Research Question 1.3

Hypothesis	Hypothesis Statements	Evaluation	Result
H 1.3: H1	**H1:** The models, when provided with a combination of Fundamental Indicators (FI) and Technical Indicators (TI) as input, yield superior forecasting accuracy for cryptocurrency closing prices compared to using either FI or TI alone with these models.	The analysis of performance metrics presented in Tables 16 to 20 and Tables 21 to 25 indicates that the combination of FI and TI consistently offers enhanced forecasting accuracy for cryptocurrency closing prices in comparison to using either FI or TI in isolation. This trend is evident across the three models: CNN-BiLSTM, BiLSTM, and ANN, with only a few exceptions noted.	Reject H0, Accept H1

Source: Author’s compilation

5.1.2.4 Main Contributions and Novelties

We made several important contributions by proposing this novel forecasting framework which can help various stakeholders including traders, short-term investors, and researchers. The main contributions can be summarized as below:

1. **Change Point Analysis**: We utilized the PELT algorithm on five cryptocurrency datasets to discern exchange price trends and identify potential non-linear patterns and change points.

2. **Multifractal Analysis**: We undertook a Multifractal Detrended Fluctuation Analysis for the cryptocurrencies, aiming to characterize their behavior daily and weekly. This approach was designed to capture the inherent complexities of cryptocurrency price movements and inform our subsequent forecasting framework.

3. **Forecasting Framework Architecture**: We proposed a forecasting framework that integrates feature selection, data preprocessing, and hybrid deep learning models. For researchers, we provide a unique dataset. For traders, we offer an accurate price forecasting framework. Our proposed financial forecasting framework has a flexible architecture applicable to cryptocurrency datasets and other financial and non-financial forecasting challenges.

4. **Signal Processing Mechanism**: We introduce a signal processing mechanism that incorporates a three-step feature selection procedure designed to refine non-linear and non-stationary time series data.

5. **Deep Learning Architecture**: We develop a hybrid deep learning architecture that incorporates a self-attention mechanism, enabling neural networks to understand the temporal aspects of non-stationary data better. We fine-tune these networks using Bayesian Optimization.

6. **Addressing Literature Gaps**: Previous literature often explores technical and fundamental indicators separately. We address this by considering their combined influence. Our forecasting model outperformed all existing studies in the literature. Moreover, we analyzed separate market phases of the cryptocurrency market and developed forecasting models for each phase for Bitcoin and Altcoins.

5.1 Conclusion and Summary of Main Findings: Objective 2

Objective 2: *To determine the price formation factors of Bitcoin and Altcoins.*

5.2.1 Recap of Objective 2

To identify the factors that impact the price dynamics of Bitcoin and Altcoins, we conducted a SHAP Analysis on the best-performing CNN-BiLSTM (TI+FI) model. This analysis encompassed all four intervals for Bitcoin and was also applied to the most recent and representative interval for the other four Altcoins.

5.2.2 Key Factors Influencing Price Formation of Bitcoin and Altcoins

5.2.2.1. Research Question 2.1

What are the key factors that influence the price formation of Bitcoin and Altcoins?

The result analysis of DeepSHAP from Figure 20 to Figure 27 reveal the key price determinants for Bitcoin, Ethereum, Ripple, Litecoin, and DASH. The major findings for price predictors are as follows:

5.2.2.1.1 Main Findings for Bitcoin’s Price Formation Factors

Interval 1: Bitcoin's Inception and Early Adoption Phase

i. **Blockchain Centric Factors**: Dominance of Blockchain network-related factors, such as the number of unspent Bitcoin transaction outputs and the cost of processing Bitcoin transactions.

ii. **Global Currency Influence**: Significant influence from global currency pairs, specifically GBP to USD and CNY to USD.
iii. **Limited Macroeconomic Impact**: Minimal influence from external macroeconomic factors, except for crude oil and global currencies to US dollar ratios.

The findings of Interval 1 are corroborated by the findings of (W. Chen et al., 2021).

Interval 2: Expansion Phase and Shift to Market-Related Factors

i. **Decline of Blockchain Factors**: Blockchain-centric factors from Interval 1 receded in their influence.
ii. **Market-Related Dominance**: Emergence of market-related factors like transaction volume, cost of transaction, and global currency ratios, especially GBP to USD and EUR to USD.
iii. **Technical Analysis**: Rise of technical indicators such as Moving Average Convergence Divergence, Volume Weighted Average Price, and Average Directional Index.

The important price determinants of this period are supported by the results of (W. Chen et al., 2021) during a similar investigation period.

Interval 3: Transitional Phase during Initial Market Peak

i. **Transactional Activity**: Fees per transaction in USD became a significant factor, reflecting heightened transactional activity.
ii. **Volume-Related Indicators**: Emergence of volume-related technical indicators like Volume Price Trend, Ease of Movement, and Accumulation/Distribution Index.
iii. **Public Interest**: Continued relevance of Google trends and tweets, indicating growing public interest.
iv. **Macroeconomic Integration**: Significance of US initial claims, NASDAQ Close, and the US 10-year treasury rate close prices, showing Bitcoin's evolving relationship with mainstream financial markets.

The important price determinants of this period are supported by the results of (W. Chen et al., 2021) during a similar investigation period.

Interval 4: Price Formation Through Multiple Market Cycles

i. **Operational Efficiency**: The most important factor was 'fees_per_transaction', suggesting that transaction costs remain a decisive factor in Bitcoin's price.
ii. **Blockchain Network Activity**: The 'confirmed_transactions_per_day' factor reinforced the importance of Blockchain network activity in Bitcoin's price determination.

iii. **Technical Indicators**: The significance of volume-centric indicators like 'volume_vpt', 'volume_em', and 'volume_adi' emphasized the role of market liquidity and trading volumes.
iv. **Public Sentiment**: Continued influence of 'google_trends' and 'Tweets', highlighting the importance of public sentiment.
v. **Macroeconomic Indicators**: The inclusion of broader macroeconomic indicators like 'US_Initial_Claims_Close', 'NASDAQ_Close', and 'US_10Y_Treasury_rate_Close' indicated a multifaceted set of factors influencing Bitcoin's price.

These findings are consistent with numerous other findings reported in past literature, as presented in Chapter 4, and discussed in detail in the results and analysis of Interval 4.

5.2.2.1.2 Main Findings for Ethereum’s Price Formation Factors

i. **Public Sentiment**: Google Search Trends for Ethereum is the most significant factor influencing Ethereum's price.
ii. **Bitcoin's Influence**: Ethereum's price is strongly influenced by Bitcoin's close price.
iii. **Blockchain Activity**: Ethereum's blockchain-specific factors, such as hash count, token transfers, and block count, play a crucial role in its price determination.
iv. **Technical Indicators**: Indicators like the Vortex Indicator Difference, Money Flow Index, and Percentage Price Oscillator Histogram are significant, suggesting reliance on technical analysis by traders.

Our results align with the observations made by (Angela & Sun, 2020), who identified a notable influence of Bitcoin prices on Ethereum values. Similarly, (H. M. Kim et al., 2021) found that the blockchain data from the Ethereum network is pivotal in shaping Ethereum's price dynamics.

5.2.2.1.3 Main Findings for Ripple’s Price Formation Factors

i. **Bitcoin's Influence**: Ripple's price is majorly influenced by the average transaction value per day in Bitcoin.
ii. **Technical Indicators**: The Aroon Indicator is significant in detecting trend changes in Ripple's price.
iii. **Social Media Sentiment**: The number of Ripple-related tweets plays a crucial role in its price formation.
iv. **Macroeconomic Elements**: Global financial dynamics, such as the Euro to US Dollar exchange rate, influence Ripple's valuation.

Our findings corroborate with the findings of (Valencia et al., 2019), who also found a strong influence of the above technical indicators and social signals on the Ripple price formation.

5.2.2.1.4 Main Findings for Price formation factors for Litecoin

i. **Social Media Sentiment**: The number of daily tweets containing the #Litecoin hashtag is the most influential factor for Litecoin's price.
ii. **Blockchain Activity**: Litecoin's blockchain variables like average block size, transaction count, and active addresses are significant in its price determination.
iii. **Bitcoin and Ethereum's Influence**: Both Bitcoin and Ethereum's close prices significantly impact Litecoin's price.
iv. **Technical Indicators**: Volatility indicators like Bollinger Bands Lower Indicator and Keltner Channel Lower Indicator highlight Litecoin's market volatility.

Our results corroborate the research of (Angela & Sun, 2020), who illustrated a marked interrelation between Litecoin values and the prices of both Bitcoin and Ethereum.

5.2.2.1.5 Main Findings for Price formation factors for DASH

i. **Bitcoin's Influence**: DASH's price is majorly influenced by Bitcoin's close price.
ii. **Blockchain Activity**: DASH's blockchain variables like transaction count, block size, and average transaction fee play a crucial role in its price formation.
iii. **Technical Indicators**: Volatility-based technical indicators, such as Volatility Bollinger Bands Percentage and Volatility Keltner Channel High, are significant for DASH's price.
iv. **Social Media Sentiment**: The number of tweets per day with #Dashpay plays a role in DASH's price formation.

5.2.3 Research Question 2.2

How does the impact of price formation factors on Bitcoin and Altcoins change across different market periods?

5.2.2.1 Main Findings for Price Formation Factors across market periods

Across the examined market intervals, the determinants influencing the price of Bitcoin and Altcoins exhibit distinct shifts. Bitcoin's initial phase was predominantly driven by its inherent blockchain metrics. As the market expanded, there was a discernible transition towards market-related factors. Transactional activity became a significant influencer by the transitional phase, with an evident correlation to specific mainstream financial market indicators. In subsequent market cycles, the emphasis was on operational efficiency, technical indicators, and select macroeconomic factors.

For Altcoins, while the influence of Bitcoin's price remains a consistent factor, each cryptocurrency displayed its unique set of determinants. Ethereum's price, for instance, is notably influenced by public sentiment, as indicated by Google trends and its blockchain metrics. Ripple's price shows sensitivity to average transaction values in Bitcoin and global financial dynamics, such as specific currency exchange rates. Litecoin's price determinants encompass social media sentiment, blockchain activity, and the influence of Bitcoin and Ethereum's closing prices. On the other hand, DASH leans towards its blockchain metrics, volatility-based technical indicators, and social media sentiment.

5.3 Conclusion and Summary of Main Findings: Objective 3

Objective 3: *To identify the factors responsible for price inconsistencies across different major cryptocurrency exchanges*

5.3.1 Recap of Objective 3

We analyzed the prices of Bitcoin, Ethereum, Ripple, Litecoin, and DASH across three prominent exchanges: Binance (Global), Kraken, and Coinbase Pro for this objective. Our objective was to identify the underlying factors of price discrepancies across these platforms. We started by analyzing the descriptive statistics of cryptocurrency prices in order to identify any anomalies. Next, we investigated absolute and relative metrics of price discrepancies, with the latter providing insight into price fluctuations in the volatile cryptocurrency landscape. After describing our model for price inconsistencies and its performance metrics, such as RMSE, MAE, and MAPE, we utilized SHAP analysis to identify the main factors contributing to these price inconsistencies.

5.3.1.1 Research Question 5.3.1

What factors contribute to price inconsistencies across major cryptocurrency exchanges (Binance, Kraken, Coinbase)?

5.3.1.1.1 Summary of Main Findings

Our analysis, as detailed in Table 39, reveals that various market, social, and technical factors contribute to price inconsistencies across the major cryptocurrency exchanges: Binance, Kraken, and Coinbase Pro. The main factors are summarized below:

1. **Kraken** (as per Figure 13 (a). Global Feature Importance for Kraken):

   i. **Bid-Ask Spread**: The difference between the highest price a buyer is willing to pay and the lowest price a seller is willing to accept is the most significant factor influencing price inconsistencies.
   ii. **Tweets**: Social media sentiment, particularly from Twitter, plays a crucial role in determining price variations.
   iii. **Maker Fees**: The fees charged for placing a new order also influence the price discrepancies on this exchange.

2. **Coinbase Pro** (as per Figure 13 (b). Global Feature Importance for Coinbase Pro):

   i. **Google Trends Data**: Public interest in cryptocurrencies, gauged by Google search volume, is the most influential factor in determining price inconsistencies.
   ii. **Williams %R**: This technical trading indicator, which shows whether a security is overbought or oversold, also plays a significant role.
   iii. **Tweets**: Like Kraken, the sentiment and discourse on Twitter have a direct impact on price variations.

3. **Binance** (as per Figure 13 (c). Global Feature Importance for Binance):

   i. **Tweets**: Given Binance's extensive following on Twitter, information or sentiment on this platform has a considerable impact on its market, making it the primary influencer of price inconsistencies.
   ii. **Bid-Ask Spread**: Market liquidity, represented by the bid-ask spread, is another key factor.
   iii. **Transaction Volume**: The volume of trades occurring on the platform also plays a significant role in determining price discrepancies.

We found that while there are common factors like tweets and the bid-ask spread influencing price inconsistencies across all three exchanges, each platform also has its unique influencers, reflecting the distinct trading dynamics and user behaviors of each exchange.

5.3.1.2 Research Question 5.3.2

How does the influence of these factors vary across different cryptocurrency exchanges (Binance, Kraken, Coinbase)?

5.3.1.2.1 Summary of Main Findings

Our results show that the influence of these factors varies across different exchanges. On Kraken, the bid-ask spread, tweets, maker fees, MACD histogram, and On-Balance Volume (OBV) are the most significant factors. For Coinbase Pro, Google trends data, Williams %R, tweets, On-Balance Volume (OBV), and the Relative Strength Index (RSI) emerge as the primary influencers. On Binance, tweets, bid-ask spread, transaction volume, MACD histogram, and the

Relative Strength Index (RSI) stand out as the key contributors. This variation suggests that each exchange has its unique dynamics and factors influencing price inconsistencies.

5.4 Conclusion and Summary of Main Findings: Objective 4

0bjective 4: *To compare the liquidity of cryptocurrency exchanges with the liquidities of different sizes of selected stocks by using Martin Liquidity Index (MLI)*

5.4.1 Recap of Objective 4

In Objective 4, we systematically compared the liquidity across major cryptocurrency exchanges and traditional stock markets. Utilizing the Martin Liquidity Index (MLI) as our primary metric, we also incorporated other measures such as Amihud's Illiquidity Ratio (AIR), AR Bid-Ask Spread, Roll's Covariance Liquidity Estimator, and the CS spread estimator. Our analysis spanned three significant cryptocurrency exchanges: Binance, Coinbase, and Kraken, juxtaposing their liquidity metrics against traditional stock markets represented by NYSE, NASDAQ, NIFTY, and BSE SENSEX.

5.4.2 Research Questions and Major Findings

5.4.2.1 Research Question 5.4.1

How does the liquidity of traditional stock exchanges (NYSE, NASDAQ, BSE Sensex, NIFTY) compare to that of major cryptocurrency exchanges (Binance, Kraken, Coin9base) as measured by the Martin Liquidity Index (MLI)?

5.4.2.1.1 Summary of Main Findings

Based on the results presented in Tables 40, 41, and 42, and visual illustrations in Figures 9 (a) to 9 (d), traditional stock exchanges, specifically NASDAQ100, NYSE Composite, and NIFTY, consistently exhibit superior liquidity compared to major cryptocurrency exchanges. NASDAQ100 and NYSE Composite emerge as the most liquid markets throughout the three intervals. Binance, as depicted in Figure 9 (a), leads among cryptocurrency exchanges in terms of the Martin Liquidity Index (MLI). Nevertheless, the analysis of Figure 9 (c) and 9 (d) reveals that Binance's liquidity improvements are insufficient to meet the benchmarks established by conventional exchanges such as NASDAQ100 and NYSE Composite, as observed in other liquidity indicators like the Average Relative Bid-Ask Spread and CS Estimator.

5.4.2.1.2 Hypothesis Evaluation

Table 46. Evaluation of Hypothesis for Research Question 5.4.1

Hypothesis	Hypothesis Statements	Evaluation	Result
H 5.4.1: H0	H0: The liquidity of major cryptocurrency exchanges (Binance, Kraken, Coinbase), as measured by the Martin Liquidity Index (MLI), is equal to or higher than that of traditional stock markets (NYSE, NASDAQ, BSE Sensex, NIFTY).	Based on Tables 40, 41, and 42, and corroborated by Figure 14 (a) and 14 (b), traditional stock exchanges consistently exhibit higher liquidity than major cryptocurrency exchanges.	Reject H0
H 5.4.1: H1	H1: The liquidity of traditional stock exchanges (NYSE, NASDAQ, BSE Sensex, NIFTY) is significantly higher than that of major cryptocurrency exchanges (Binance, Kraken, Coinbase) as measured by the Martin Liquidity Index (MLI).	The data from Tables 40, 41, and 42, along with Figure 14 (a) and 14 (b), confirm that traditional stock exchanges have superior liquidity compared to major cryptocurrency exchanges.	Accept H1

Source: Author’s compilation

5.4.3 Research Questions and Major Findings

5.4.3.1 Research Question 2

Which of the major cryptocurrency exchanges (Binance, Kraken, Coinbase) exhibits the highest liquidity as measured by the Martin Liquidity Index (MLI)?

5.4.3.1.1 Summary of Main Findings

According to the findings shown in Tables 40, 41, and 42, along with the visual representations provided in Figures 9 (a) to 9 (d), Binance consistently demonstrates the highest level of liquidity across cryptocurrency exchanges, as evaluated by the Martin Liquidity Index (MLI). Among the three major cryptocurrency exchanges, Binance ranks as the best in liquidity, although Kraken and Coinbase exhibit significant levels of liquidity.

5.5 Implications of this Research

The implications of our findings are manifold and extend to various stakeholders, including researchers, traders, investors, and policymakers.

For Objective 1, the proposed flexible architecture, which includes data preparation, feature selection, data preprocessing, model training, hyperparameter optimization, model performance evaluation, price prediction, and model explainability, provides a comprehensive framework for financial forecasting studies. This framework can be applied to various financial assets beyond cryptocurrencies, offering researchers and financial analysts’ valuable insights. The framework applies signal processing methods to pre-process the non-linear financial data. A novel three-step feature selection process is proposed, integrating fundamental and technical indicators, and using advanced machine learning models contributes to both the technical and fundamental analysis literature in cryptocurrencies.

For Objective 2, the SHAP analysis results highlight the complex role of various factors in cryptocurrency price formation in different market phases, including transactional attributes of the Blockchain network, market liquidity conditions, public sentiment, and broader economic indicators. This understanding can guide investors and traders in anticipating price movements and developing effective trading strategies.

For Objective 3, identifying factors contributing to price inconsistencies across significant cryptocurrency exchanges can inform high-frequency traders and arbitrageurs looking to exploit these price discrepancies for profit. Moreover, the significance of exchange-specific features highlights the idiosyncratic factors driving the pricing in each given exchange.

For Objective 4, comparing liquidity across traditional stock exchanges and major cryptocurrency exchanges using the Martin Liquidity Index (MLI) provides valuable insights for various market participants. For traders and investors, understanding the liquidity of an exchange can inform decisions related to transaction costs, price slippage, and the speed of trade execution. For policymakers and regulators, these findings can inform decisions about market structure and investor protection measures.

5.6 Limitations of this Research and Future Research Directions

While our analysis stands rigorous in its approach, certain limitations warrant attention. Firstly, our predictive models, though effective in current conditions, are grounded in indicators tailored to the prevailing dynamics of the cryptocurrency markets. As the nature of these markets is inherently evolving, novel indicators may surface in the future. These new metrics could offer more granular insights or capture emergent patterns, potentially enhancing the predictive accuracy beyond the capabilities of our current models.

Secondly, while our study primarily concentrated on established cryptocurrencies and their associated exchanges, it is important to recognize that our findings may not seamlessly extend to smaller or less established platforms. Future research could broaden the scope of this analysis, exploring the behavior and dynamics of nascent or lesser-known exchanges and their associated cryptocurrencies.

Third, while our research provides a comprehensive liquidity analysis using multiple recognized metrics and primarily leverages OHLCV data, it is essential to recognize the depth and granularity that order book data can offer. At its core, liquidity is significantly influenced by the microstructures present in the bid-ask spread and the depth of buy and sell orders at different price levels. By utilizing only OHLCV data, we operate at a broader level, which may miss out on nuanced, intra-day fluctuations and strategic order placements. Therefore, a potential avenue for future research lies in harvesting extensive order book data at lower frequencies, such as 5-minute or hourly intervals, over an extended period, like a decade. This approach would allow researchers to delve deeper into the liquidity landscape of cryptocurrencies, understanding the complexities within the bid-ask spread. Future studies can juxtapose the findings from such detailed order book analysis against the liquidity indicators we employed in this study. The comparison can shed light on each method's advantages and potential shortcomings, ensuring that future liquidity assessments are rigorous and holistic. However, liquidity is a multifaceted concept, and there is yet to be a universally agreed-upon measure to quantify it. Therefore, future studies could consider evaluating other approaches to measure liquidity.

Finally, while this study provides comprehensive insights into the pricing behavior of Bitcoin and major Altcoins, it is important to recognize the dynamic nature of the cryptocurrency market. With its rapid evolution, new cryptocurrencies, potentially rising to prominence, could surface post this research timeline. This is not a limitation but rather a characteristic of this ever-evolving domain. Future studies might extend the focus to encompass the behavior of emerging major cryptocurrencies, particularly emphasizing the growing segment of stablecoins such as Tether (USDT), USD Coin (USDC), and DAI. This will further enrich our understanding of pricing dynamics and the overarching influence on the cryptocurrency ecosystem.

References

1. Abdi, F., & Ranaldo, A. (2017). A simple estimation of bid-Ask spreads from daily close, high, and low prices. *Review of Financial Studies*, *30*(12), 4437–4480. https://doi.org/10.1093/rfs/hhx084
2. Abhyankar, A. A., & Singla, H. K. (2022). Comparing predictive performance of general regression neural network (GRNN) and hedonic regression model for factors affecting housing prices in “Pune-India.” *International Journal of Housing Markets and Analysis*, *15*(2), 451–477. https://doi.org/10.1108/IJHMA-01-2021-0003
3. Aggarwal, D., Chandrasekaran, S., & Annamalai, B. (2020). A complete empirical ensemble mode decomposition and support vector machine-based approach to predict Bitcoin prices. *Journal of Behavioral and Experimental Finance*, *27*, 100335. https://doi.org/10.1016/j.jbef.2020.100335
4. Agur, I., Ari, A., & Dell’Ariccia, G. (2022). Designing central bank digital currencies. *Journal of Monetary Economics*, *125*, 62–79. https://doi.org/10.1016/j.jmoneco.2021.05.002
5. Agyei, S. K., Adam, A. M., Bossman, A., Asiamah, O., Owusu Junior, P., Asafo-Adjei, R., & Asafo-Adjei, E. (2022). Does volatility in cryptocurrencies drive the interconnectedness between the cryptocurrencies market? Insights from wavelets. *Cogent Economics and Finance*, *10*(1). https://doi.org/10.1080/23322039.2022.2061682
6. Akba, F., Medeni, I. T., Guzel, M. S., & Askerzade, I. (2021). Manipulator detection in cryptocurrency markets based on forecasting anomalies. *IEEE Access*, *9*, 108819–108831. https://doi.org/10.1109/ACCESS.2021.3101528
7. Akcora, C. G., Gel, Y. R., & Kantarcioglu, M. (2022). Blockchain networks: Data structures of Bitcoin, Monero, Zcash, Ethereum, Ripple, and Iota. *WIREs Data Mining and Knowledge Discovery*, *12*(1). https://doi.org/10.1002/widm.1436
8. Alexander, C., & Heck, D. (2020). Price Discovery in Bitcoin: The Impact of Unregulated Markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3583843
9. Ali, O., Ally, M., Clutterbuck, & Dwivedi, Y. (2020). The state of play of blockchain technology in the financial services sector: A systematic literature review. In *International Journal of Information Management* (Vol. 54). https://doi.org/10.1016/j.ijinfomgt.2020.102199
10. Ali, O., Momin, M., Shrestha, A., Das, R., Alhajj, F., & Dwivedi, Y. K. (2023). A review of the key challenges of non-fungible tokens. *Technological Forecasting and Social Change*, *187*, 122248. https://doi.org/10.1016/j.techfore.2022.122248
11. Aljadani, A. (2022). DLCP2F: a DL-based cryptocurrency price prediction framework. *Discover Artificial Intelligence*, *2*(1), 20. https://doi.org/10.1007/s44163-022-00036-2
12. Aloosh, A., & Ouzan, S. (2020). The psychology of cryptocurrency prices. *Finance Research Letters*, *33*, 101192. https://doi.org/10.1016/j.frl.2019.05.010
13. Altan, A., Karasu, S., & Bekiros, S. (2019). Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques. *Chaos, Solitons and Fractals*, *126*, 325–336. https://doi.org/10.1016/j.chaos.2019.07.011
14. Al-Yahyaee, K. H., Mensi, W., Ko, H. U., Yoon, S. M., & Kang, S. H. (2020). Why cryptocurrency markets are inefficient: The impact of liquidity and volatility. *North American Journal of Economics and Finance*, *52*, 101168. https://doi.org/10.1016/j.najef.2020.101168
15. Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6
16. Ammous, S. (2015). Economics beyond Financial Intermediation: Digital Currencies’ Possibilities for Growth, Poverty Alleviation, and International Development. *Journal of Private Enterprise*, *30*(3), 19–50.
17. Andrade, D. M. de, Barros, F., Motoki, F. Y., & Oliveira da Silva, M. (2021). Price dynamics of cryptocurrencies in parallel markets: evidence from Bitcoin exchanges in Brazil. *Studies in Economics and Finance*, *38*(5), 1040–1053. https://doi.org/10.1108/SEF-11-2020-0450
18. Angela, O., & Sun, Y. (2020). Factors affecting cryptocurrency prices: Evidence from ethereum. *Proceedings of 2020 International Conference on Information Management and Technology, ICIMTech 2020*, *August*, 318–323. https://doi.org/10.1109/ICIMTech50083.2020.9211195
19. Auer, B. R., & Rottmann, H. (2019). Have capital market anomalies worldwide attenuated in the recent era of high liquidity and trading activity? *Journal of Economics and Business*, *103*, 61–79. https://doi.org/10.1016/j.jeconbus.2018.12.003

20. Bakshy, E., Dworkin, L., Karrer, B., Kashin, K., Letham, B., Murthy, A., & Facebook, S. S. (2018). AE: A domain-agnostic platform for adaptive experimentation. In *NIPS* (Issue Nips).
21. Balcilar, M., Bouri, E., Gupta, R., & Roubaud, D. (2017a). Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling*, *64*(March), 74–81. https://doi.org/10.1016/j.econmod.2017.03.019
22. Balcilar, M., Bouri, E., Gupta, R., & Roubaud, D. (2017b). Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling*, *64*, 74–81. https://doi.org/10.1016/j.econmod.2017.03.019
23. Bartoletti, M., Lande, S., Loddo, A., Pompianu, L., & Serusi, S. (2021). Cryptocurrency scams: Analysis and perspectives. *IEEE Access*, *9*, 148353–148373. https://doi.org/10.1109/ACCESS.2021.3123894
24. Baur, D. G., Hong, K. H., & Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money*, *54*, 177–189. https://doi.org/10.1016/j.intfin.2017.12.004
25. Benlagha, N., & Hemrit, W. (2023). Asymmetric determinants of Bitcoin’s wild price movements. *Managerial Finance*, *49*(2), 227–247. https://doi.org/10.1108/MF-03-2022-0105
26. Bianchetti, M., Ricci, C., & Scaringi, M. (2017). Are Cryptocurrencies Real Financial Bubbles? Evidence from Quantitative Analyses. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3092427
27. Biryukov, A., & Tikhomirov, S. (2019). Security and privacy of mobile wallet users in Bitcoin, Dash, Monero, and Zcash. *Pervasive and Mobile Computing*, *59*, 101030. https://doi.org/10.1016/j.pmcj.2019.101030
28. Blau, B. M. (2018). Price dynamics and speculative trading in Bitcoin. *Research in International Business and Finance*, *43*, 15–21. https://doi.org/10.1016/j.ribaf.2017.07.183
29. Borri, N., & Shakhnov, K. (2020). Regulation spillovers across cryptocurrency markets. *Finance Research Letters*, *36*, 101333. https://doi.org/10.1016/j.frl.2019.101333
30. Bouoiyour, J., & Selmi, R. (2017). *The Bitcoin price formation: Beyond the fundamental sources*. *June*. https://doi.org/10.13140/RG.2.2.23880.32000
31. Brauneis, A., Mestel, R., Riordan, R., & Theissen, E. (2018). A High-Frequency Analysis of Bitcoin Markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3249477
32. Brauneis, A., Mestel, R., Riordan, R., & Theissen, E. (2021a). How to measure the liquidity of cryptocurrency markets? *Journal of Banking and Finance*, *124*, 106041. https://doi.org/10.1016/j.jbankfin.2020.106041
33. Brauneis, A., Mestel, R., Riordan, R., & Theissen, E. (2021b). How to measure the liquidity of cryptocurrency markets? *Journal of Banking and Finance*, *124*, 106041. https://doi.org/10.1016/j.jbankfin.2020.106041
34. Brauneis, A., Mestel, R., Riordan, R., & Theissen, E. (2022). Bitcoin unchained: Determinants of cryptocurrency exchange liquidity. *Journal of Empirical Finance*, *69*, 106–122. https://doi.org/10.1016/j.jempfin.2022.08.004
35. Caporale, G. M., Gil-Alana, L., & Plastun, A. (2018). Persistence in the cryptocurrency market. *Research in International Business and Finance*, *46*(December 2017), 141–148. https://doi.org/10.1016/j.ribaf.2018.01.002
36. Carbó, J. M., & Gorjon, S. (2022a). Application of Machine Learning Models and Interpretability Techniques to Identify the Determinants of the Price of Bitcoin. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4087481
37. Carbó, J. M., & Gorjon, S. (2022b). Application of Machine Learning Models and Interpretability Techniques to Identify the Determinants of the Price of Bitcoin. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4087481
38. Cermak, V. (2017). Can Bitcoin Become a Viable Alternative to Fiat Currencies? An Empirical Analysis of Bitcoin’s Volatility Based on a GARCH Model. *SSRN Electronic Journal*, *May*. https://doi.org/10.2139/ssrn.2961405
39. Chen, W., Xu, H., Jia, L., & Gao, Y. (2021). Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. *International Journal of Forecasting*, *37*(1), 28–43. https://doi.org/10.1016/j.ijforecast.2020.02.008
40. Chen, Y. (2018). Blockchain tokens and the potential democratization of entrepreneurship and innovation. *Business Horizons*, *61*(4), 567–575. https://doi.org/10.1016/j.bushor.2018.03.006
41. Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics*, *365*, 112395. https://doi.org/10.1016/j.cam.2019.112395
42. Cheung, A. (Wai-K., Roca, E., & Su, J.-J. (2015). Crypto-currency bubbles: an application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox bitcoin prices. *Applied Economics*, *47*(23), 2348–2358. https://doi.org/10.1080/00036846.2015.1005827

43. Ciaian, P., Rajcaniova, M., & Kancs, d'Artis. (2016). The economics of BitCoin price formation. *Applied Economics*, *48*(19), 1799–1815. https://doi.org/10.1080/00036846.2015.1109038
44. Ciaian, P., Rajcaniova, M., & Kancs, d'Artis. (2018). Virtual relationships: Short- and long-run evidence from BitCoin and altcoin markets. *Journal of International Financial Markets, Institutions and Money*, *52*, 173–195. https://doi.org/10.1016/j.intfin.2017.11.001
45. Coen, A., & de La Bruslerie, H. (2019). The Informational Dimensions of the Amihud (2002) Illiquidity Measure. *Finance Research Letters*, *29*, 23–29. https://doi.org/10.1016/j.frl.2019.03.015
46. Cohen, G. (2021). Trading cryptocurrencies using second order stochastic dominance. *Mathematics*, *9*(22). https://doi.org/10.3390/math9222861
47. Corbet, S., Larkin, C., Lucey, B. M., Meegan, A., & Yarovaya, L. (2020). The impact of macroeconomic news on Bitcoin returns. *European Journal of Finance*, *26*(14), 1396–1416. https://doi.org/10.1080/1351847X.2020.1737168
48. 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. https://doi.org/10.1016/j.irfa.2018.09.003
49. Corbet, S., Meegan, A., Larkin, C. J., Lucey, B. M., & Yarovaya, L. (2017). Exploring the Dynamic Relationships between Cryptocurrencies and Other Financial Assets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3070288
50. Cortez, K., Rodríguez-García, M. D. P., & Mongrut, S. (2021). Exchange market liquidity prediction with the k-nearest neighbor approach: Crypto vs. fiat currencies. *Mathematics*, *9*(1), 1–15. https://doi.org/10.3390/math9010056
51. Corwin, S. A., & Schultz, P. (2012). A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices. *Journal of Finance*, *67*(2), 719–760. https://doi.org/10.1111/j.1540-6261.2012.01729.x
52. Darolles, S., Le Fol, G., & Mero, G. (2011). When Market Illiquidity Generates Volume. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1102089
53. Díaz, A., & Escribano, A. (2020). Measuring the multi-faceted dimension of liquidity in financial markets: A literature review. *Research in International Business and Finance*, *51*, 101079. https://doi.org/10.1016/j.ribaf.2019.101079
54. Dimpfl, T., & Peter, F. J. (2021). Nothing but noise? Price discovery across cryptocurrency exchanges. *Journal of Financial Markets*, *54*, 100584. https://doi.org/10.1016/j.finmar.2020.100584
55. Disli, M., Abd Rabbo, F., Leneeuw, T., & Nagayev, R. (2022). Cryptocurrency comovements and crypto exchange movement: The relocation of Binance. *Finance Research Letters*, *48*, 102989. https://doi.org/10.1016/j.frl.2022.102989
56. Dorcas Wambui, G. (2015). The Power of the Pruned Exact Linear Time(PELT) Test in Multiple Changepoint Detection. *American Journal of Theoretical and Applied Statistics*, *4*(6), 581. https://doi.org/10.11648/j.ajtas.20150406.30
57. Doumenis, Y., Izadi, J., Dhamdhere, P., Katsikas, E., & Koufopoulos, D. (2021). A Critical Analysis of Volatility Surprise in Bitcoin Cryptocurrency and Other Financial Assets. *Risks*, *9*(11), 207. https://doi.org/10.3390/risks9110207
58. Dwyer, G. P. (2015). The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*, *17*, 81–91. https://doi.org/10.1016/j.jfs.2014.11.006
59. El-Berawi, A. S., Belal, M. A. F., & Ellatif, M. M. A. (2021). Adaptive Deep Learning based Cryptocurrency Price Fluctuation Classification. *International Journal of Advanced Computer Science and Applications*, *12*(12), 487–500. https://doi.org/10.14569/IJACSA.2021.0121264
60. Eom, C., Kaizoji, T., Hoon, S., & Pichl, L. (2019). Bitcoin and investor sentiment : Statistical characteristics and predictability. *Physica A*, *514*, 511–521. https://doi.org/10.1016/j.physa.2018.09.063
61. Feinstein, B. D., & Werbach, K. (2020). The Impact of Cryptocurrency Regulation on Trading Markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3649475
62. Ferdiansyah, Othman, S. H., Radzi, R. Z. M., Stiawan, D., & Sutikno, T. (2023). Hybrid gated recurrent unit bidirectional-long short-term memory model to improve cryptocurrency prediction accuracy. *IAES International Journal of Artificial Intelligence*, *12*(1), 251–261. https://doi.org/10.11591/ijai.v12.i1.pp251-261
63. Field, J., & Inci, A. C. (2023). Risk translation: how cryptocurrency impacts company risk, beta and returns. *Journal of Capital Markets Studies*, *7*(1), 5–21. https://doi.org/10.1108/JCMS-02-2023-0003
64. Fleischer, J. P., von Laszewski, G., Theran, C., & Bautista, Y. J. P. (2022). Time Series Analysis of Cryptocurrency Prices Using Long Short-Term Memory. *Algorithms*, *15*(7), 1–13. https://doi.org/10.3390/a15070230

65. Gaies, B., Nakhli, M. S., Sahut, J.-M., & Schweizer, D. (2023). Interactions between investors' fear and greed sentiment and Bitcoin prices. *The North American Journal of Economics and Finance*, *67*, 101924. https://doi.org/10.1016/j.najef.2023.101924
66. Gao, W., & Su, C. (2020). Analysis of earnings forecast of blockchain financial products based on particle swarm optimization. *Journal of Computational and Applied Mathematics*, *372*, 112724. https://doi.org/10.1016/j.cam.2020.112724
67. García-Medina, A., & Huynh, T. L. D. (2021). What drives bitcoin? An approach from continuous local transfer entropy and deep learning classification models. *Entropy*, *23*(12), 1–19. https://doi.org/10.3390/e23121582
68. Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D. N., & Giaglis, G. M. (2015). Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2607167
69. Ghasemiyeh, R., Moghdani, R., & Sana, S. S. (2017). A Hybrid Artificial Neural Network with Metaheuristic Algorithms for Predicting Stock Price. *Cybernetics and Systems*, *48*(4), 365–392. https://doi.org/10.1080/01969722.2017.1285162
70. Giudici, P., & Pagnottoni, P. (2019). High frequency price change spillovers in bitcoin markets. *Risks*, *7*(4), 111. https://doi.org/10.3390/risks7040111
71. Goczek, Ł., & Skliarov, I. (2019). What drives the Bitcoin price? A factor augmented error correction mechanism investigation. *Applied Economics*, *51*(59), 6393–6410. https://doi.org/10.1080/00036846.2019.1619021
72. Gozbasi, O., Altinoz, B., & Sahin, E. E. (2021). IS BITCOIN A SAFE HAVEN? A STUDY ON THE FACTORS THAT AFFECT BITCOIN PRICES. *International Journal of Economics and Financial Issues*, *11*(4), 35–40. https://doi.org/10.32479/ijefi.11602
73. Gyamerah, S. A. (2020). On Forecasting the Intraday Bitcoin Price using Ensemble of Variational Mode Decomposition and Generalized Additive Model. *Journal of King Saud University - Computer and Information Sciences*, *xxxx*. https://doi.org/10.1016/j.jksuci.2020.01.006
74. H. Sabah. (2023). Prospective Empirical Study on the Determinants of Bitcoin Price Formation (Case Study on Morocco). *Journal of Applied Business and Economics*, *25*(2). https://doi.org/10.33423/jabe.v25i2.6096
75. Hakim das Neves, R. (2020). Bitcoin pricing: impact of attractiveness variables. *Financial Innovation*, *6*(1), 21. https://doi.org/10.1186/s40854-020-00176-3
76. Hamayel, M. J., & Owda, A. Y. (2021). A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. *Ai*, *2*(4), 477–496. https://doi.org/10.3390/ai2040030
77. Hammami, H., & Boujelbene, Y. (2022). The effects of stock market crises shocks on market liquidity, market volatility and exchange rate volatility: Case of the Tunisian stock market. *International Journal of Finance*, *7*(1), 40–58. https://doi.org/10.47941/ijf.828
78. Härdle, W. K., Harvey, C. R., & Reule, R. C. G. (2020). Understanding Cryptocurrencies. In *Journal of Financial Econometrics* (Vol. 18, Issue 2, pp. 181–208). https://doi.org/10.1093/jjfinec/nbz033
79. Hayes, A. S. (2017). Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telematics and Informatics*, *34*(7), 1308–1321. https://doi.org/10.1016/j.tele.2016.05.005
80. Hu, Y., Ni, J., & Wen, L. (2020). A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction. *Physica A: Statistical Mechanics and Its Applications*, *557*, 124907. https://doi.org/10.1016/j.physa.2020.124907
81. Huang, J. Z., Huang, W., & Ni, J. (2019). Predicting bitcoin returns using high-dimensional technical indicators. *Journal of Finance and Data Science*, *5*(3), 140–155. https://doi.org/10.1016/j.jfds.2018.10.001
82. Ishikawa, M. (2017). Designing Virtual Currency Regulation in Japan: Lessons from the Mt Gox Case. *Journal of Financial Regulation*, fjw015. https://doi.org/10.1093/jfr/fjw015
83. Islam, M. S., Hussain, I., Rahman, M. M., Park, S. J., & Hossain, M. A. (2022). Explainable Artificial Intelligence Model for Stroke Prediction Using EEG Signal. *Sensors (Basel, Switzerland)*, *22*(24). https://doi.org/10.3390/s22249859
84. James, N., & Menzies, M. (2022a). Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time. *Nonlinear Dynamics*, *107*(4), 4001–4017. https://doi.org/10.1007/s11071-021-07166-9
85. James, N., & Menzies, M. (2022b). Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time. *Nonlinear Dynamics*, *107*(4), 4001–4017. https://doi.org/10.1007/s11071-021-07166-9

86. Jang, H., & Lee, J. (2018a). An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information. *IEEE Access*, *6*, 5427–5437. https://doi.org/10.1109/ACCESS.2017.2779181
87. Jang, H., & Lee, J. (2018b). An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information. *IEEE Access*, *6*, 5427–5437. https://doi.org/10.1109/ACCESS.2017.2779181
88. Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, *6*(3), 255–259. https://doi.org/10.1016/0165-1765(80)90024-5
89. Jha, A. K., & Jha, R. (2020). India's Response to COVID-19 Crisis. *The Indian Economic Journal*, *68*(3), 341–351. https://doi.org/10.1177/0019466220976685
90. Ji, Q., Bouri, E., Keung, C., Lau, M., & Roubaud, D. (2018). International Review of Financial Analysis Dynamic connectedness and integration in cryptocurrency markets. *International Review of Financial Analysis*, *December*, 1–16. https://doi.org/10.1016/j.irfa.2018.12.002
91. Jouini, E. (2003). Market Imperfections, Equilibrium and Arbitrage. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1007202
92. Kavianpour, P., Kavianpour, M., Jahani, E., & Ramezani, A. (2021). *A CNN-BiLSTM Model with Attention Mechanism for Earthquake Prediction*. 1–13. http://arxiv.org/abs/2112.13444
93. Kawai, D., Cuevas, A., Routledge, B., Soska, K., Zetlin-Jones, A., & Christin, N. (2023). Is your digital neighbor a reliable investment advisor? *Proceedings of the ACM Web Conference 2023*, 3581–3591. https://doi.org/10.1145/3543507.3583502
94. KAYAHAN, C., GÖKGÖZ, H., & MURAT, T. (2022). QUANTILE CONNECTEDNESS BETWEEN CRYPTOCURRENCIES AND STABLECOINS. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, *37*, 143–156. https://doi.org/10.18092/ulikidince.1146239
95. Kayal, P., & Rohilla, P. (2021). Bitcoin in the economics and finance literature: a survey. *SN Business & Economics*, *1*(7). https://doi.org/10.1007/s43546-021-00090-5
96. Kerr, D. S., Loveland, K. A., Smith, K. T., & Smith, L. M. (2023). Cryptocurrency Risks, Fraud Cases, and Financial Performance. *Risks*, *11*(3), 51. https://doi.org/10.3390/risks11030051
97. Kim, G., Shin, D.-H., Choi, J. G., & Lim, S. (2022). A Deep Learning-Based Cryptocurrency Price Prediction Model That Uses On-Chain Data. *IEEE Access*, *10*, 56232–56248. https://doi.org/10.1109/access.2022.3177888
98. Kim, H. M., Bock, G. W., & Lee, G. (2021). Predicting Ethereum prices with machine learning based on Blockchain information. *Expert Systems with Applications*, *184*(February 2020), 115480. https://doi.org/10.1016/j.eswa.2021.115480
99. Koutmos, D., & Wei, W. C. (2023). Nowcasting bitcoin's crash risk with order imbalance. *Review of Quantitative Finance and Accounting*. https://doi.org/10.1007/s11156-023-01148-1
100. Kristoufek, L. (2015). What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis. *PLoS ONE*, *10*(4), 1–15. https://doi.org/10.1371/journal.pone.0123923
101. Kukacka, J., & Kristoufek, L. (2023). Fundamental and speculative components of the cryptocurrency pricing dynamics. *Financial Innovation*, *9*(1), 61. https://doi.org/10.1186/s40854-023-00465-7
102. Kumar, D., Meghwani, S. S., & Thakur, M. (2016). Proximal support vector machine based hybrid prediction models for trend forecasting in financial markets. *Journal of Computational Science*, *17*, 1–13. https://doi.org/10.1016/j.jocs.2016.07.006
103. Kumar, P., Alok, S., & Pandey, K. (2022). A new grey system approach to forecast closing price of Bitcoin, Bionic, Cardano, Dogecoin, Ethereum, XRP Cryptocurrencies. *Quality & Quantity*, *0123456789*. https://doi.org/10.1007/s11135-022-01463-0
104. Kumar, S., Liu, J., & Scutella, J. (2015). The impact of supply chain disruptions on stockholder wealth in India. *International Journal of Physical Distribution and Logistics Management*, *45*(9–10), 938–958. https://doi.org/10.1108/IJPDLM-09-2013-0247
105. Kuo Chuen, D. L., Guo, L., & Wang, Y. (2017). Cryptocurrency: A New Investment Opportunity? *The Journal of Alternative Investments*, *20*(3), 16–40. https://doi.org/10.3905/jai.2018.20.3.016
106. Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the boruta package. *Journal of Statistical Software*, *36*(11), 1–13. https://doi.org/10.18637/jss.v036.i11
107. Lahmiri, S., & Bekiros, S. (2018). Chaos, randomness and multi-fractality in Bitcoin market. *Chaos, Solitons and Fractals*, *106*, 28–34. https://doi.org/10.1016/j.chaos.2017.11.005
108. Lahmiri, S., & Bekiros, S. (2019a). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons and Fractals*, *118*, 35–40. https://doi.org/10.1016/j.chaos.2018.11.014

109. Lahmiri, S., & Bekiros, S. (2019b). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons and Fractals*, *118*, 35–40. https://doi.org/10.1016/j.chaos.2018.11.014
110. Lahmiri, S., & Bekiros, S. (2020). Intelligent forecasting with machine learning trading systems in chaotic intraday Bitcoin market. *Chaos, Solitons and Fractals*, *133*. https://doi.org/10.1016/j.chaos.2020.109641
111. Lamothe-Fernández, P., Alaminos, D., Lamothe-López, P., & Fernández-Gámez, M. A. (2020). Deep learning methods for modeling bitcoin price. *Mathematics*, *8*(8), 1–13. https://doi.org/10.3390/MATH8081245
112. Ledolter, J. (2013). Local Polynomial Regression: A Nonparametric Regression Approach. In *Data Mining and Business Analytics with R* (pp. 55–66). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118596289.ch4
113. Lee, A. D., Li, M., & Zheng, H. (2020). Bitcoin: Speculative asset or innovative technology? *Journal of International Financial Markets, Institutions and Money*, *67*. https://doi.org/10.1016/j.intfin.2020.101209
114. Lee, C. F., & Lee, J. C. (2015). Methods of Denoising Financial Data. *Handbook of Financial Econometrics and Statistics*, *August 2015*, v–vi. https://doi.org/10.1007/978-1-4614-7750-1
115. Leirvik, T. (2022). Cryptocurrency returns and the volatility of liquidity. *Finance Research Letters*, *44*, 102031. https://doi.org/10.1016/j.frl.2021.102031
116. Li, X., & Wang, C. A. (2017). The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin. *Decision Support Systems*, *95*, 49–60. https://doi.org/10.1016/j.dss.2016.12.001
117. Liang, C., Zhang, Y., Li, X., & Ma, F. (2020). Which predictor is more predictive for Bitcoin volatility? And why? *International Journal of Finance and Economics*, *July*, 1–15. https://doi.org/10.1002/ijfe.2252
118. Liew, J. K.-S., & Hewlett, L. (2017). The Case for Bitcoin for Institutional Investors: Bubble Investing or Fundamentally Sound? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3082808
119. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. *2008 Eighth IEEE International Conference on Data Mining*, 413–422. https://doi.org/10.1109/ICDM.2008.17
120. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2012). Isolation-Based Anomaly Detection. *ACM Transactions on Knowledge Discovery from Data*, *6*(1), 1–39. https://doi.org/10.1145/2133360.2133363
121. Liu, L. (2019). Are Bitcon returns predictable?: Evidence from technical indicators. *Physica A: Statistical Mechanics and Its Applications*, *533*, 121950. https://doi.org/10.1016/j.physa.2019.121950
122. Livieris, I. E., Kiriakidou, N., Stavroyiannis, S., & Pintelas, P. (2021). An advanced CNN-LSTM model for cryptocurrency forecasting. *Electronics (Switzerland)*, *10*(3), 1–16. https://doi.org/10.3390/electronics10030287
123. Loi, H. (2017). The Liquidity of Bitcoin. *International Journal of Economics and Finance*, *10*(1), 13. https://doi.org/10.5539/ijef.v10n1p13
124. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *2017-Decem*(Section 2), 4766–4775.
125. Luo, C., Pan, L., Chen, B., & Xu, H. (2022). Bitcoin Price Forecasting: An Integrated Approach Using Hybrid LSTM-ELM Models. *Mathematical Problems in Engineering*, *2022*(Dl). https://doi.org/10.1155/2022/2126518
126. Lybek, T., & Sarr, A. (2002). Measuring Liquidity in Financial Markets. *IMF Working Papers*, *02*(232), 1. https://doi.org/10.5089/9781451875577.001
127. Maciel, L., Ballini, R., Gomide, F., & Yager, R. (2022a). Forecasting cryptocurrencies prices using data driven level set fuzzy models. *Expert Systems with Applications*, *210*(August), 118387. https://doi.org/10.1016/j.eswa.2022.118387
128. Maciel, L., Ballini, R., Gomide, F., & Yager, R. (2022b). Forecasting cryptocurrencies prices using data driven level set fuzzy models. *Expert Systems with Applications*, *210*(July), 118387. https://doi.org/10.1016/j.eswa.2022.118387
129. Makarov, I., & Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics*, *135*(2), 293–319. https://doi.org/10.1016/j.jfineco.2019.07.001
130. Mallqui, D. C. A., & Fernandes, R. A. S. (2019). Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. *Applied Soft Computing Journal*, *75*, 596–606. https://doi.org/10.1016/j.asoc.2018.11.038
131. Massaoudi, M., Refaat, S. S., Abu-Rub, H., Chihi, I., & Oueslati, F. S. (2020). PLS-CNN-BiLSTM: An end-to-end algorithm-based savitzky-golay smoothing and evolution strategy for load forecasting. *Energies*, *13*(20), 1–29. https://doi.org/10.3390/en13205464
132. Mayo, D., & Elgazzar, H. (2022). Predicting Cryptocurrency Price Change Direction from Supply-Side Factors via Machine Learning Methods. *2022 IEEE World AI IoT Congress, AIIoT 2022*, 330–336. https://doi.org/10.1109/AIIoT54504.2022.9817249

133. Meynkhard, A. (2020). *Effect of Bitcoin Volatility on Altcoins Pricing* (pp. 652–664). https://doi.org/10.1007/978-3-030-63322-6_55
134. Miljkovic, D. (1999). The Law of One Price in International Trade: A Critical Review. *Applied Economic Perspectives and Policy*, *21*(1), 126–139. https://doi.org/10.2307/1349976
135. Miller, D., & Kim, J.-M. (2021). Univariate and Multivariate Machine Learning Forecasting Models on the Price Returns of Cryptocurrencies. *Journal of Risk and Financial Management*, *14*(10), 486. https://doi.org/10.3390/jrfm14100486
136. Miloş, L. R., Haţiegan, C., Miloş, M. C., Barna, F. M., & Boțoc, C. (2020). Multifractal Detrended Fluctuation Analysis (MF-DFA) of Stock Market Indexes. Empirical Evidence from Seven Central and Eastern European Markets. *Sustainability*, *12*(2), 535. https://doi.org/10.3390/su12020535
137. Misra, N., Rao, T. J., Gupta, S., & Grima, L. (2023). Blockchain technology for the financial markets. In *Intelligent Multimedia Technologies for Financial Risk Management: Trends, tools and applications* (pp. 225–260). Institution of Engineering and Technology. https://doi.org/10.1049/PBPC060E_ch11
138. Mohanty, D. (2018). Ethereum Use Cases. In *Ethereum for Architects and Developers* (pp. 203–243). Apress. https://doi.org/10.1007/978-1-4842-4075-5_9
139. Morillon, T. G., & Chacon, R. G. (2022). Dissecting the stock to flow model for Bitcoin. *Studies in Economics and Finance*, *39*(3), 506–523. https://doi.org/10.1108/SEF-10-2021-0409
140. Mudassir, M., Bennbaia, S., Unal, D., & Hammoudeh, M. (2020). Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. *Neural Computing and Applications*, *6*. https://doi.org/10.1007/s00521-020-05129-6
141. Munim, Z. H., Shakil, M. H., & Alon, I. (2019). Next-Day Bitcoin Price Forecast. *Journal of Risk and Financial Management*, *12*(2), 103. https://doi.org/10.3390/jrfm12020103
142. Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. *Www.Bitcoin.Org*, 9. https://doi.org/10.1007/s10838-008-9062-0
143. Ning, Y., Ong, M. E. H., Chakraborty, B., Goldstein, B. A., Ting, D. S. W., Vaughan, R., & Liu, N. (2022). Shapley variable importance cloud for interpretable machine learning. *Patterns*, *3*(4), 100452. https://doi.org/10.1016/j.patter.2022.100452
144. Nosratabadi, S., Mosavi, A., Duan, P., Ghamisi, P., Filip, F., Band, S. S., Reuter, U., Gama, J., & Gandomi, A. H. (2020). Data science in economics: Comprehensive review of advanced machine learning and deep learning methods. *Mathematics, 8*(10), 1–25. https://doi.org/10.3390/math8101799
145. Ntakaris, A., Kanniainen, J., Gabbouj, M., & Iosifidis, A. (2020). Mid-price prediction based on machine learning methods with technical and quantitative indicators. *PLoS ONE*, *15*(6 June), 1–39. https://doi.org/10.1371/journal.pone.0234107
146. Ong, B., Lee, T. M., Li, G., & Chuen, D. L. K. (2015). Evaluating the Potential of Alternative Cryptocurrencies. In *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data* (pp. 81–135). https://doi.org/10.1016/B978-0-12-802117-0.00005-9
147. Ortu, M., Uras, N., Conversano, C., Bartolucci, S., & Destefanis, G. (2022). On technical trading and social media indicators for cryptocurrency price classification through deep learning. *Expert Systems with Applications, 198*, 116804. https://doi.org/10.1016/j.eswa.2022.116804
148. Pagnottoni, P., & Dimpfl, T. (2019). Price discovery on Bitcoin markets. *Digital Finance*, *1*(1–4), 139–161. https://doi.org/10.1007/s42521-019-00006-x
149. Parekh, R., Patel, N. P., Thakkar, N., Gupta, R., Tanwar, S., Sharma, G., Davidson, I. E., & Sharma, R. (2022). DL-GuesS: Deep Learning and Sentiment Analysis-Based Cryptocurrency Price Prediction. *IEEE Access*, *10*(April 2021), 35398–35409. https://doi.org/10.1109/ACCESS.2022.3163305
150. Park, M., & Chai, S. (2020). The effect of information asymmetry on investment behavior in cryptocurrency market. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2020-Janua*, 4043–4052. https://doi.org/10.24251/hicss.2020.494
151. Patel, N. P., Parekh, R., Thakkar, N., Gupta, R., Tanwar, S., Sharma, G., Davidson, I. E., & Sharma, R. (2022). Fusion in Cryptocurrency Price Prediction: A Decade Survey on Recent Advancements, Architecture, and Potential Future Directions. *IEEE Access*, *10*, 34511–34538. https://doi.org/10.1109/ACCESS.2022.3163023
152. Pichl, L., & Kaizoji, T. (2017). Volatility Analysis of Bitcoin Price Time Series. *Quantitative Finance and Economics*, *1*(4), 474–485. https://doi.org/10.3934/QFE.2017.4.474
153. Pieters, G., & Vivanco, S. (2017). Financial regulations and price inconsistencies across Bitcoin markets. *Information Economics and Policy*, *39*(2015), 1–14. https://doi.org/10.1016/j.infoecopol.2017.02.002

154. Radojičić, D., Radojičić, N., & Kredatus, S. (2021). A multicriteria optimization approach for the stock market feature selection. *Computer Science and Information Systems*, *18*(3), 749–769. https://doi.org/10.2298/CSIS200326044R
155. Ranjan, S., Kayal, P., & Saraf, M. (2023). Bitcoin Price Prediction: A Machine Learning Sample Dimension Approach. *Computational Economics*, *61*(4), 1617–1636. https://doi.org/10.1007/s10614-022-10262-6
156. Ranjit, M. P., Ganapathy, G., Sridhar, K., & Arumugham, V. (2019). Efficient deep learning hyperparameter tuning using cloud infrastructure: Intelligent distributed hyperparameter tuning with Bayesian optimization in the cloud. *IEEE International Conference on Cloud Computing, CLOUD*, *2019-July*, 520–522. https://doi.org/10.1109/CLOUD.2019.00097
157. Rashid, N. A., & Ismail, M. T. (2023). Modelling and Forecasting the Trend in Cryptocurrency Prices. *Journal of Information and Communication Technology*, *22*(3), 449–501. https://doi.org/10.32890/jict2023.22.3.6
158. Rathee, N., Ankita Singh, Sharda, T., Goel, N., Mansi Aggarwal, & Dudeja, S. (2023). Analysis and price prediction of cryptocurrencies for historical and live data using ensemble-based neural networks. *Knowledge and Information Systems*. https://doi.org/10.1007/s10115-023-01871-0
159. ROLL, R. (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance*, *39*(4), 1127–1139. https://doi.org/10.1111/j.1540-6261.1984.tb03897.x
160. Saad, M., Choi, J., Nyang, D., Kim, J., & Mohaisen, A. (2020). Toward characterizing blockchain-based cryptocurrencies for highly accurate predictions. *IEEE Systems Journal*, *14*(1), 321–332. https://doi.org/10.1109/JSYST.2019.2927707
161. Salman, A., & Razzaq, M. G. A. (2018). Bitcoin and the World of Digital Currencies. In *Financial Management from an Emerging Market Perspective*. InTech. https://doi.org/10.5772/intechopen.71294
162. Sauer, B. (2016). Virtual Currencies, the Money Market, and Monetary Policy. *International Advances in Economic Research*, *22*(2), 117–130. https://doi.org/10.1007/s11294-016-9576-x
163. Schafer, R. W. (2011). What Is a Savitzky-Golay Filter? [Lecture Notes]. *IEEE Signal Processing Magazine*, *28*(4), 111–117. https://doi.org/10.1109/MSP.2011.941097
164. Schellinger, B. (2020). Optimization of special cryptocurrency portfolios. *Journal of Risk Finance*, *21*(2), 127–157. https://doi.org/10.1108/JRF-11-2019-0221
165. Schmid, M., Rath, D., & Diebold, U. (2022). Why and How Savitzky–Golay Filters Should Be Replaced. *ACS Measurement Science Au*, *2*(2), 185–196. https://doi.org/10.1021/acsmeasuresciau.1c00054
166. Schueffel, P. (2021). DeFi: Decentralized Finance - An Introduction and Overview. *Journal of Innovation Management*, *9*(3), I–XI. https://doi.org/10.24840/2183-0606_009.003_0001
167. Sensoy, A. (2019). The inefficiency of Bitcoin revisited: A high-frequency analysis with alternative currencies. *Finance Research Letters*, *28*, 68–73. https://doi.org/10.1016/j.frl.2018.04.002
168. Seo, J., Ma, H., & Saha, T. K. (2018). On savitzky-golay filtering for online condition monitoring of transformer on-load tap changer. *IEEE Transactions on Power Delivery*, *33*(4), 1689–1698. https://doi.org/10.1109/TPWRD.2017.2749374
169. Shahbazi, Z., & Byun, Y. C. (2021). Improving the cryptocurrency price prediction performance based on reinforcement learning. *IEEE Access*, *9*, 162651–162659. https://doi.org/10.1109/ACCESS.2021.3133937
170. Shi, S., Li, J., Li, G., Pan, P., & Liu, K. (2021). XPM: An Explainable Deep Reinforcement Learning Framework for Portfolio Management. *International Conference on Information and Knowledge Management, Proceedings*, 1661–1670. https://doi.org/10.1145/3459637.3482494
171. Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning important features through propagating activation differences. *34th International Conference on Machine Learning, ICML 2017*, *7*, 4844–4866.
172. Shynkevich, A. (2023). Law of one price and return on Arbitrage Trading: Bitcoin vs. Ethereum. *Journal of Economics and Finance*, *47*(3), 763–792. https://doi.org/10.1007/s12197-023-09631-0
173. Soni, T. K. (2012). Effectiveness of Artificial Neural Networks in Forecasting BSE Sensex Index Values. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1863187
174. Stancel, D. (2015). *Economic consequences of cryptocurrencies and associated decentralized systems*. *May*, 1–45.
175. Sugunnasil, P., & Somhom, S. (2010). *Feature Selection for Neural Network Based Stock Prediction*. 137–138.
176. Szetela, B., Mentel, G., & Gędek, S. (2016). Dependency Analysis between Bitcoin and Selected Global Currencies. *Dynamic Econometric Models*, *16*(1 SE-Articles), 133–144. https://doi.org/10.12775/DEM.2016.009

177. Takaishi, T. (2018). Statistical properties and multifractality of Bitcoin. *Physica A: Statistical Mechanics and Its Applications*, *506*, 507–519. https://doi.org/10.1016/j.physa.2018.04.046
178. Takaishi, T., & Adachi, T. (2020). Market Efficiency, Liquidity, and Multifractality of Bitcoin: A Dynamic Study. *Asia-Pacific Financial Markets*, *27*(1), 145–154. https://doi.org/10.1007/s10690-019-09286-0
179. Tang, Q., Fan, T., Shi, R., Huang, J., & Ma, Y. (2021). *Prediction of financial time series using LSTM and data denoising methods*.
180. Tanwar, S., Patel, N. P., Patel, S. N., Patel, J. R., Sharma, G., & Davidson, I. E. (2021a). Deep Learning-Based Cryptocurrency Price Prediction Scheme with Inter-Dependent Relations. *IEEE Access*, *9*, 138633–138646. https://doi.org/10.1109/ACCESS.2021.3117848
181. Tanwar, S., Patel, N. P., Patel, S. N., Patel, J. R., Sharma, G., & Davidson, I. E. (2021b). Deep Learning-Based Cryptocurrency Price Prediction Scheme with Inter-Dependent Relations. *IEEE Access*, *9*, 138633–138646. https://doi.org/10.1109/ACCESS.2021.3117848
182. Tartakovsky, E. (2020). Auto-Correlation of Returns in Major Cryptocurrency Markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3639430
183. Tripathi, B., & Sharma, R. K. (2022). Modeling Bitcoin Prices using Signal Processing Methods, Bayesian Optimization, and Deep Neural Networks. *Computational Economics*, *0123456789*. https://doi.org/10.1007/s10614-022-10325-8
184. Troster, V., Tiwari, A. K., Shahbaz, M., & Macedo, D. N. (2019). Bitcoin returns and risk: A general GARCH and GAS analysis. *Finance Research Letters*, *30*, 187–193. https://doi.org/10.1016/j.frl.2018.09.014
185. Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. In *Signal Processing* (Vol. 167). https://doi.org/10.1016/j.sigpro.2019.107299
186. Tsang, K. P., & Yang, Z. (2020). Price dispersion in bitcoin exchanges. *Economics Letters*, *194*, 109379. https://doi.org/10.1016/j.econlet.2020.109379
187. Valencia, F., Gómez-Espinosa, A., & Valdés-Aguirre, B. (2019). Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *Entropy*, *21*(6). https://doi.org/10.3390/e21060589
188. Wątorek, M., Drożdż, S., Kwapień, J., Minati, L., Oświęcimka, P., & Stanuszek, M. (2021). Multiscale characteristics of the emerging global cryptocurrency market. *Physics Reports*, *901*, 1–82. https://doi.org/10.1016/j.physrep.2020.10.005
189. Wątorek, M., Kwapień, J., & Drożdż, S. (2023). Cryptocurrencies Are Becoming Part of the World Global Financial Market. In *Entropy* (Vol. 25, Issue 2). https://doi.org/10.3390/e25020377
190. Wei, W. C. (2018). Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, *168*, 21–24. https://doi.org/10.1016/j.econlet.2018.04.003
191. Woebbeking, F. (2021). Cryptocurrency volatility markets. *Digital Finance*, *3*(3–4), 273–298. https://doi.org/10.1007/s42521-021-00037-3
192. Xin, B., & Peng, W. (2020). Prediction for Chaotic Time Series-Based AE-CNN and Transfer Learning. *Complexity*, *2020*, 1–9. https://doi.org/10.1155/2020/2680480
193. Yamada, M. (2022). Profitability and liquidity provision of HFTs during large price shocks: Does relative tick size matter? *Finance Research Letters*, *46*, 102308. https://doi.org/10.1016/j.frl.2021.102308
194. Yang, Y., Xiong, J., Zhao, L., Wang, X., Hua, L., & Wu, L. (2023). A Novel Method of Blockchain Cryptocurrency Price Prediction Using Fractional Grey Model. *Fractal and Fractional*, *7*(7), 547. https://doi.org/10.3390/fractalfract7070547
195. Ye, Z., Wu, Y., Chen, H., Pan, Y., & Jiang, Q. (2022). A Stacking Ensemble Deep Learning Model for Bitcoin Price Prediction Using Twitter Comments on Bitcoin. *Mathematics*, *10*(8). https://doi.org/10.3390/math10081307
196. Yermack, D. (2015). Is Bitcoin a Real Currency? An Economic Appraisal. In *Handbook of Digital Currency* (pp. 31–43). Elsevier. https://doi.org/10.1016/B978-0-12-802117-0.00002-3
197. Younis Masiha, R. (2022). Effects of Cryptocurrencies on Global Economics: A Review Study. *Qubahan Academic Journal*, *2*(2), 9–15. https://doi.org/10.48161/qaj.v2n2a126
198. Yue, W., Zhang, S., & Zhang, Q. (2021). Asymmetric News Effects on Cryptocurrency Liquidity: an Event Study Perspective. *Finance Research Letters*, *41*, 101799. https://doi.org/10.1016/j.frl.2020.101799
199. Zhang, W., Wang, P., Li, X., & Shen, D. (2018). The inefficiency of cryptocurrency and its cross-correlation with Dow Jones Industrial Average. *Physica A: Statistical Mechanics and Its Applications*, *510*, 658–670. https://doi.org/10.1016/j.physa.2018.07.032
200. Zhu, Y., Dickinson, D., & Li, J. (2017). Erratum to: Analysis on the influence factors of Bitcoin's price based on VEC model. *Financial Innovation*, *3*(1), 7. https://doi.org/10.1186/s40854-017-0057-x

201. Zhu, Y., Ma, J., Gu, F., Wang, J., Li, Z., Zhang, Y., Xu, J., Li, Y., Wang, Y., & Yang, X. (2023). Price Prediction of Bitcoin Based on Adaptive Feature Selection and Model Optimization. *Mathematics*, *11*(6), 1335. https://doi.org/10.3390/math11061335
202. Zulfiqar, M., Gamage, K. A. A., Kamran, M., & Rasheed, M. B. (2022). Hyperparameter Optimization of Bayesian Neural Network Using Bayesian Optimization and Intelligent Feature Engineering for Load Forecasting. *Sensors*, *22*(12). https://doi.org/10.3390/s22124446

Appendix A

Table A1. Optimal Hyperparameter Configurations Recommended by Bayesian Optimization for Bitcoin (BTC) for Interval 4

Hyperparameter Name	**ANN (FI)**	**ANN (FI+TI)**	**ANN (TI)**	**BiLSTM (FI)**	**BiLSTM (FI+TI)**	**BiLSTM (TI)**	**CNN-BiLSTM (FI)**	**CNN-BiLSTM (FI+TI)**	**CNN-BiLSTM (TI)**
Learning rate	0.0001	0.0015	0.0019	0.0071	0.0013	0.0042	0.0068	0.0074	0.007
Dropout rate	0.125	0.2	0.2	0.2	0.2	0.11	0.15	0.2	0.15
Count of hidden layers	3	4	3	3	3	2	3	3	3
Neuron count per layer	459	480	470	450	460	455	465	475	460
Batch size	32	16	32	64	16	64	16	8	16
Activation function	relu	elu	sigmoid	tanh	relu	tanh	elu	relu	sigmoid
Optimizer	Adamx	rms	Adam	Nadam	rms	Adamx	sgd	Adam	rms
CNN filters	-	-	-	-	-	-	128	64	128
Kernel size	-	-	-	-	-	-	3	4	3
Pooling size and type	-	-	-	-	-	-	2(max)	3(average)	2 (max)
Gradient Clipping	5.2	4.8	5	4.9	5.1	5	4.7	4.6	4.8
Weight Decay (L2 reg.)	0.0023	0.0018	0.002	0.0019	0.0021	0.0022	0.0012	0.0016	0.0027

Source: Author's Compilation.

Table A2. Optimal Hyperparameter Configurations Recommended by Bayesian Optimization for Ethereum (ETH) for Interval 4

Hyperparameter Name	**ANN (FI)**	**ANN (FI+TI)**	**ANN (TI)**	**BiLSTM (FI)**	**BiLSTM (FI+TI)**	**BiLSTM (TI)**	**CNN-BiLSTM (FI)**	**CNN-BiLSTM (FI+TI)**	**CNN-BiLSTM (TI)**
Learning rate	0.0001	0.0012	0.0011	0.007	0.0014	0.0045	0.0065	0.0072	0.0069
Dropout rate	0.2	0.2	0.2	0.2	0.25	0.2	0.25	0.2	0.2
Count of hidden layers	3	3	3	3	3	3	3	3	3
Neuron count per layer	457	478	468	448	458	444	463	450	458
	32	16	32	8	16	8	16	8	16
Activation function	relu	elu	sigmoid	tanh	relu	tanh	elu	relu	sigmoid
Optimizer	Adamx	rms	Adam	Nadam	rms	Adamx	sgd	Adamx	rms
CNN filters	-	-	-	-	-	-	128	64	128
Kernel size	-	-	-	-	-	-	3	4	3
Pooling size and type	-	-	-	-	-	-	2(max)	2(average)	2(max)
Gradient Clipping	5.1	4.7	4.9	4.8	5	4.9	4.6	4.5	4.7
Weight Decay (L2 reg.)	0.0022	0.0017	0.0019	0.0018	0.002	0.0021	0.0016	0.0015	0.0016

Source: Author's Compilation.

Table A3. Optimal Hyperparameter Configurations Recommended by Bayesian Optimization for Ripple (XRP) for Interval 4

Hyperparameter Name	**ANN (FI)**	**ANN (FI+TI)**	**ANN (TI)**	**BiLSTM (FI)**	**BiLSTM (FI+TI)**	**BiLSTM (TI)**	**CNN-BiLSTM (FI)**	**CNN-BiLSTM (FI+TI)**	**CNN-BiLSTM (TI)**
Learning rate	0.0009	0.0013	0.001	0.0065	0.0012	0.0043	0.0062	0.007	0.0067
Dropout rate	0.12	0.2	0.18	0.2	0.1	0.11	0.12	0.2	0.2
Count of hidden layers	3	3	3	3	2	3	3	3	3
Neuron count per layer	487	476	464	445	455	455	460	473	480
Batch size	32	16	8	8	16	32	32	8	16
Activation function	relu	elu		tanh	relu	tanh	elu	relu	sigmoid

			sigmoid						
Optimizer	Adamx	rms	Adam	Nadam	rms	Adamx	sgd	Adamx	rms
CNN filters	-	-	-	-	-	-	128	64	128
Kernel size	-	-	-	-	-	-	3	4	3
Pooling size and type	-	-	-	-	-	-	2	2(average)	2 (max)
Gradient Clipping	5	4.6	4.8	4.7	4.9	4.8	4.5	4.4	4.6
Weight Decay (L2 reg.)	0.002	0.0016	0.0018	0.0017	0.0019	0.002	0.0015	0.0014	0.0015

Source: Author’s Compilation.

Table A4. Optimal Hyperparameter Configurations Recommended by Bayesian Optimization for Litecoin for Interval 4

Hyperparameter Name	ANN (FI)	ANN (FI+TI)	ANN (TI)	BiLSTM (FI)	BiLSTM (FI+TI)	BiLSTM (TI)	CNN-BiLSTM (FI)	CNN-BiLSTM (FI+TI)	CNN-BiLSTM (TI)
Learning rate	0.0008	0.0014	0.0011	0.0063	0.0011	0.0041	0.006	0.0069	0.0065
Dropout rate	0.25	0.2	0.2	0.2	0.2	0.2	0.18	0.25	0.2
Count of hidden layers	2	3	3	2	3	3	3	3	3
Neuron count per layer	450	478	465	440	425	445	419	465	450
Batch size	32	16	32	8	8	8	16	8	16
Activation function	relu	elu	sigmoid	tanh	relu	tanh	elu	relu	sigmoid
Optimizer	Adamx	rms	Adam	Nadam	Adamx	Adamx	sgd	Adamx	rms
CNN filters	-	-	-	-	-	-	120	60	120
Kernel size	-	-	-	-	-	-	3	4	3
Pooling size and type	-	-	-	-	-	-	2(max)	3(average)	2(max)
Gradient Clipping	4.9	4.5	4.7	4.6	4.8	4.7	4.4	4.3	4.5
Weight Decay (L2 reg.)	0.0029	0.0015	0.0017	0.0016	0.0018	0.0019	0.0014	0.0013	0.0014

Source: Author’s Compilation.

Table A5. Optimal Hyperparameter Configurations Recommended by Bayesian Optimization for DASH for Interval 4

Hyperparameter Name	ANN (FI)	ANN (FI+TI)	ANN (TI)	BiLSTM (FI)	BiLSTM (FI+TI)	BiLSTM (TI)	CNN-BiLSTM (FI)	CNN-BiLSTM (FI+TI)	CNN-BiLSTM (TI)
Learning rate	0.0002	0.0013	0.0012	0.0069	0.0015	0.0043	0.0064	0.0071	0.0068
Dropout rate	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Count of hidden layers	3	3	3	3	3	3	3	3	3
Neuron count per layer	455	476	466	446	456	442	461	448	456
Batch size	32	32	16	64	32	64	32	16	32
Activation function	relu	elu	sigmoid	tanh	relu	tanh	relu	sigmoid	relu
Optimizer	Adam	rms	Adam	Nadam	rms	Adamx	sgd	Adamx	rms
CNN filters	-	-	-	-	-	-	128	128	64
Kernel size	-	-	-	-	-	-	3	3	4
Pooling size and type	-	-	-	-	-	-	2(max)	3(max)	2(average)
Gradient Clipping	5	4.6	4.8	4.7	4.9	4.8	4.5	4.4	4.6
Weight Decay (L2 reg.)	0.0021	0.0016	0.0018	0.0017	0.0019	0.002	0.0015	0.0014	0.0015

Source: Author’s Compilation.

Table A6. Mann-Whitney U Test Results for Bitcoin (BTC)

Change Point Segment 1	Change Point Segment 2	U-Statistic	P-Value
220 to 330	330 to 395	0	**2.50093E-28**
330 to 395	395 to 685	18673	**3.98491E-35**
395 to 685	685 to 865	51850	**2.36414E-72**
685 to 865	865 to 1405	296	**7.0104E-89**
865 to 1405	1405 to 1460	3	**2.3141E-34**
1405 to 1460	1460 to 1505	0	**1.02933E-17**
1460 to 1505	1505 to 1605	5	**8.60467E-22**
1505 to 1605	1605 to 1685	8000	**1.12636E-30**
1605 to 1685	1685 to 1745	13	**9.28253E-24**
1685 to 1745	1745 to 1805	0	**3.55601E-21**
1745 to 1805	1805 to 1835	1800	**1.37091E-14**
1805 to 1835	1835 to 1960	3706	**1.12743E-16**
1835 to 1960	1960 to 1995	4375	**1.76644E-19**
1960 to 1995	1995 to 2220	7875	**1.85512E-21**
1995 to 2220	2220 to 2291	1267	**1.16926E-26**

Source: Author’s Calculation. The values in **bold**, indicate significant p-values.

Table A7. Mann-Whitney U Test Results for Ethereum (ETH)

Change Point Segment 1	Change Point Segment 2	U-Statistic	P-Value
140 to 345	345 to 435	0	**1.45574E-42**
345 to 435	435 to 585	13124	**1.8925E-34**
435 to 585	585 to 1310	108748	**5.62649E-83**
585 to 1310	1310 to 1470	108	**4.29578E-87**
1310 to 1470	1470 to 1555	0	**6.02762E-38**
1470 to 1555	1555 to 1685	104	**5.45727E-34**
1555 to 1685	1685 to 1755	923	**1.56039E-20**
1685 to 1755	1755 to 1835	102	**2.93652E-24**
1755 to 1835	1835 to 1960	9988	**2.25406E-33**
1835 to 1960	1960 to 1995	4369	**2.2142E-19**
1960 to 1995	1995 to 2291	9984	**2.90616E-19**

Source: Author’s Calculation. The values in **bold**, indicate significant p-values.

Table A8. Mann-Whitney U Test Results for Dash (DASH)

Change Point Segment 1	Change Point Segment 2	U-Statistic	P-Value
70 to 150	150 to 230	10	**1.36578E-27**
150 to 230	230 to 315	0	**1.49977E-28**
230 to 315	315 to 395	13	**2.40607E-28**
315 to 395	395 to 430	2249	**2.50345E-07**
395 to 430	430 to 525	3293	**1.1732E-17**
430 to 525	525 to 580	5207	**4.66914E-24**
525 to 580	580 to 690	6040	**2.00626E-25**
580 to 690	690 to 820	14299	**1.3696E-40**
690 to 820	820 to 950	38	**9.11008E-44**
820 to 950	950 to 1425	60413	**8.37894E-63**
950 to 1425	1425 to 1510	3385	**2.15874E-34**
1425 to 1510	1510 to 1565	0	**2.03657E-23**
1510 to 1565	1565 to 1605	96	**3.90361E-14**
1565 to 1605	1605 to 1690	3400	**2.36861E-19**
1605 to 1690	1690 to 1805	1179	**5.0116E-20**
1690 to 1805	1805 to 1960	17818	**9.46434E-45**
1805 to 1960	1960 to 2291	51305	**1.05826E-70**

Source: Author’s Calculation. The values in **bold**, indicate significant p-values.

Table A9. Mann-Whitney U Test Results for Litecoin (LTC)

Change Point Segment 1	Change Point Segment 2	U-Statistic	P-Value
125 to 235	235 to 345	216	**4.43664E-35**
235 to 345	345 to 380	0	**5.99886E-19**
345 to 380	380 to 450	2422	**4.17723E-16**
380 to 450	450 to 530	5272	**1.27331E-20**

450 to 530	530 to 585	4399	**7.19401E-23**
530 to 585	585 to 685	5500	**8.44655E-25**
585 to 685	685 to 800	11272	**6.81554E-34**
685 to 800	800 to 875	8	**3.40673E-31**
800 to 875	875 to 925	11	**5.93446E-21**
875 to 925	925 to 1000	3715	**1.85782E-20**
925 to 1000	1000 to 1420	30588	**1.17099E-38**
1000 to 1420	1420 to 1460	162	**1.13888E-24**
1420 to 1460	1460 to 1505	3	**2.94435E-15**
1460 to 1505	1505 to 1565	15	**5.56116E-18**
1505 to 1565	1565 to 1605	27	**1.58785E-16**
1565 to 1605	1605 to 1635	1200	**1.12111E-12**
1605 to 1635	1635 to 1685	1491	**1.85088E-13**
1635 to 1685	1685 to 1760	44	**2.83913E-20**
1685 to 1760	1760 to 1805	161	**1.31608E-16**
1760 to 1805	1805 to 1850	2025	**3.17369E-16**
1805 to 1850	1850 to 1960	4892	**1.63393E-21**
1850 to 1960	1960 to 2160	22000	**4.5502E-48**
1960 to 2160	2160 to 2291	504	**1.60082E-49**

Source: Author's Calculation. The values in **bold**, indicate significant p-values.

Table A10. Mann-Whitney U Test Results for Ripple (XRP)

Change Point Segment 1	Change Point Segment 2	U-Statistic	P-Value
125 to 345	345 to 395	65	**1.10802E-27**
345 to 395	395 to 435	1594	**1.44129E-06**
395 to 435	435 to 530	3597	**2.96051E-16**
435 to 530	530 to 585	5078	**7.02027E-22**
530 to 585	585 to 630	2458	**2.84324E-17**
585 to 630	630 to 695	14	**1.33109E-18**
630 to 695	695 to 955	16501	**1.48173E-32**
695 to 955	955 to 1425	120703	**8.5892E-106**
955 to 1425	1425 to 1455	0	**4.00248E-20**
1425 to 1455	1455 to 1495	1193	**2.04059E-12**
1455 to 1495	1495 to 1560	56	**2.2989E-16**
1495 to 1560	1560 to 1605	0	**6.20985E-19**
1560 to 1605	1605 to 1635	1273	**1.03451E-10**
1605 to 1635	1635 to 1690	1617	**3.37198E-13**
1635 to 1690	1690 to 1805	21	**1.28584E-25**
1690 to 1805	1805 to 1960	17768	**2.86766E-44**
1805 to 1960	1960 to 2291	51305	**1.05826E-70**

Source: Author's Calculation. The values in **bold**, indicate significant p-values.

Appendix B

B1.1 List of Publications and Patents

Table B1. List of publications

S No.	Title of the paper; Name of Journal, Year, and DOI.	Authors	Impact Factor in SSCI and SCIE list	Proof of journal being in SSCI/SCIE
1	Modeling Bitcoin Prices using Signal Processing Methods, Bayesian Optimization, and Deep Neural Networks; Computational Economics; (2022); https://doi.org/10.1007/s10614-022-10325-8	Bhaskar Tripathi; Dr. Rakesh Kumar Sharma	Impact Factor: 1.741 (2021); 5-Year Impact Factor: 1.720 (2021)	Click here
2	EEG-Based Emotion Classification in Financial Trading Using Deep Learning: Effects of Risk Control Measures; Sensors (Basel, Switzerland), 23(7); DOI: https://doi.org/10.3390/s23073474	Bhaskar Tripathi; Dr. Rakesh Kumar Sharma	Impact Factor: 3.847 (2021); 5-Year Impact Factor: 4.050 (2021)	Click here

Table B2. List of patents

S No.	Title, Publication date, Patent office, Case number,	Inventors	Patent Type	Link to Patent Grant
1	A mind-controlled portfolio optimization and backtesting system for online trading; 10/13/2022, Germany (DE), 20 2022 103 759.0	Bhaskar Tripathi; Dr. Rakesh Kumar Sharma	Utility Patent	Click here
2	A blockchain and IoT based system to improve network security; 09/29/2022, Germany (DE), 20 2022 104 472.4	Bhaskar Tripathi; Dr. Rakesh Kumar Sharma	Utility Patent	Click here

Table B2. Conference Paper

S No.	Title, Volume, Issue, Link	Authors	ISSN
1	Empirical Analysis of Blockchain Based Security for the Internet of Things Using Structural Equation Model Approach; Gurugram University Business Review (GUBR), Volume 2, Issue 2, July–December 2022, https://gurugramuniversity.ac.in/gujbr/GUBR%20Vol%202%20Issue%202%20JULY-DEC%202022.pdf#page=103	Bhaskar Tripathi; Dr. Rakesh Kumar Sharma	ISSN: 2582-9718 (Online)