Bitcoin Price Forecasting: An Analysis of Diverse Models

Faraz Mustafa

Dissertation submitted to International Business School
for the partial fulfilment of the requirement for the degree of
MASTER OF SCIENCE IN IT FOR BUSINESS DATA ANALYTICS
May and 2025

Declaration

I hereby declare that this Capstone Project report, titled "Bitcoin Price Prediction Using Sentiment Analysis and Deep Learning," is my original work, conducted under academic supervision. All sources used are appropriately cited, and the project adheres to ethical research standards.

Acknowledgment

I express my gratitude to my supervisor for their guidance, my peers for their feedback, and my family for their support throughout this project. Their encouragement was invaluable.

Executive Summary

Bitcoin's inherent volatility presents significant challenges and opportunities. Predicting its price accurately is complex, often beyond traditional forecasting methods. This project aimed to address this by developing advanced predictive models using diverse data and sophisticated machine learning to capture the cryptocurrency market's dynamic nature.

The research employed a multi-faceted methodology. Data collection included historical price and volume for Bitcoin and five other major cryptocurrencies (Ethereum, Binance Coin, Cardano, Solana, XRP) from 2020 to 2025. Bitcoin-related X (Twitter) posts from 2021 to 2023 were also collected for sentiment analysis using a BERT model. Data preprocessing involved cleaning tweet text and handling price data. Exploratory Data Analysis (EDA) provided insights into market trends, interdependencies, and sentiment patterns. Feature engineering generated predictors like lagged prices, technical indicators, temporal features, and multi-cryptocurrency data. Random Forest feature importance identified key predictors, leading to the exclusion of sentiment data from final models due to limited impact.

The project implemented and compared six models: ARIMA, Prophet, Random Forest, XGBoost, LSTM, and GRU. Each was trained and evaluated on an unseen test set using MAE, RMSE, MAPE, R^2, and Directional Accuracy.

Results showed a clear performance hierarchy. Traditional models (ARIMA, Prophet) performed poorly on the test set. Ensemble methods (Random Forest, XGBoost) showed significant overfitting, with strong validation performance but poor test results. Deep learning models, however, performed significantly better. The LSTM was the top performer.

Key findings indicate that models capable of handling sequential data and non-linearities (LSTM, GRU) are best for volatile cryptocurrency markets. Multi-cryptocurrency data was beneficial for capturing market influences. Limitations include challenges in predicting extreme price swings and computational constraints.

The research has significant implications. The validated LSTM and GRU models offer a robust foundation for developing accurate prediction tools for traders, investors, and financial institutions, aiding risk management and decision-making. Future work should address predicting extreme events, explore more data sources (macroeconomic, on-chain), and investigate advanced architectures. This project provides valuable insights and a practical framework for Bitcoin price prediction in a data-driven financial landscape.

Table of Contents

1. Introduction

Bitcoin was launched in 2009 by Satoshi Nakamoto. This has revolutionized the financial landscape as the first decentralized cryptocurrency. It operates on a blockchain to facilitate peer-to-peer transactions without intermediaries (Satoshi Nakamoto, 2008; Digiconomist, n.d.). The significance of bitcoin lies in its ability to provide financial autonomy while also enabling users to bypass traditional banking systems. Banking institutions often impose high fees and certain restriction on money transfer (OSL, 2025). By May 2025, Bitcoin's market capitalization exceeded $1.2 trillion (Economic Times, 2025). Hence, making it a very important asset for investors, traders, and institutions globally. However, it is a highly volatile asset which is often traded on the basis of speculation. These speculations are based on the current market information and past trends (Pepperstone, 2024). Traders often keep in mind the regulatory announcements, politcal sutiuations and macroeconomic shifts (Mudrex, 2025; OSL, 2025). Hence, predicting the prices of bitcoin accurately becomes very difficult and is often considered a gamble.

This project focuses on addressing the difficulty in forecasting Bitcoin prices due to their non-linear nature. We will see in this project that the traditional financial models like ARIMA struggle to capture trends accurately. These models assume linear relationships and stationarity, these assumption don't usually stand in the case of bitcoin. An example in this case is that of the 50% drop in May 2021 following China's regulatory crackdown (Aidoo & Ababio, 2023; Hua, 2020). This is a prime example of where the prices fluctuate and cause volatility which is very difficult for any model to predict. Machine learning, particularly deep learning models like Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU) offers a promising solution. Machine learning models capture temporal dependencies and non-linear patterns (Modi et al., 2023; Souza & Silva, 2024). Additionally, sentiment analysis of social media platforms like X can capture real-time market sentiment, which behavioral finance suggests drives price movements through investor psychology.

The objective of this project is to develop a predictive model for Bitcoin prices by integrating sentiment analysis of X posts along side prices of other selected crypto currencies with historical price data, using LSTM, GRU, Random Forest, Prophet, XG Boost and ARIMA models. We will see whether the sentiment analysis has a direct relationship with the movements in the price of bitcoin. We will also look at whether sentiment scores will play an important role in the model training or not.

Specific goals include collecting and preprocessing X posts and cryptocurrency price data from 2020 to 2023. We will generate sentiment scores, conducting exploratory data analysis to identify trends and correlations.

We will be implementing LSTM, GRU, XG Boost, Random forest and linear regression models. Then we will be evaluating their performance using metrics like Mean Squared Error (MSE) and R^2. The study also aims to provide actionable business insights for traders and financial institutions, addressing gaps in existing literature by incorporating multi-cryptocurrency data and sentiment-driven forecasting.

The scope of the project focuses on Bitcoin as the primary asset, with price and volume data from five other cryptocurrencies (Ethereum, Binance Coin, Cardano, Solana, XRP) to capture market interdependencies. The sentiment analysis is limited to English-language X posts from February 2021 to January 2023, due to data availability. The price data spans January 2020 to May 2025 to cover multiple market cycles. We will use feature importance to check which features are important for model training. We will be training multiple models to address the lack of data and check effects of our features to help us test our hypothesises.

1.1 Business and Technical Significance

The business significance of this project lies in its potential to provide traders and investors with tools to navigate Bitcoin's volatility, improving decision-making and returns. Accurate predictions can inform trading strategies, risk management, and portfolio allocation. For financial institutions, sentiment-based models offer a competitive edge in market analysis. Technically, the project advances data analytics by integrating NLP with deep learning, specifically Long Short-Term Memory (LSTM), GRU, XG Boost, Random forest, Prophet and ARIMA, to model complex temporal relationships. This interdisciplinary approach bridges finance and technology, contributing to the evolving field of financial forecasting.

1.2 Problem Statement

The high volatility of Bitcoin prices complicates investment decisions. Existing models often overlook social media sentiment, a potential driver of market movements. This project addresses the question: Can sentiment analysis of X posts, integrated with machine learning models, accurately predict Bitcoin price changes?

Secondly, we will also look at whether the prices and volume traded of other cyptos also play a role in accurately predicting the price of Bitcoin.

1.3 Hypothesis Statements

Hypothesis Statement 1:

Ho: There is no statistically significant relationship between sentiment scores derived from X and the price of Bitcoin.

H1: There is a statistically significant relationship between sentiment scores derived from X and the price of Bitcoin.

Hypothesis Statement 2:

Ho: The prices and trading volumes of Ethereum (ETH), Binance Coin (BNB), Cardano (ADA), Solana (SOL), and Ripple (XRP) do not collectively have a statistically significant impact on the prediction of Bitcoin prices.

H1: The prices and trading volumes of Ethereum (ETH), Binance Coin (BNB), Cardano (ADA), Solana (SOL), and Ripple (XRP) collectively have a statistically significant impact on the prediction of Bitcoin prices.

1.4 Objectives

1. Collect and preprocess X posts and historical cryptocurrency price data.
2. Perform sentiment analysis to generate sentiment scores.
3. Develop LSTM, GRU, XG Boost, Random forest, Prophet and ARIMA models to predict Bitcoin price changes.
4. Evaluate model performance against baseline methods.
5. Drive actionable insights for stakeholders.

1.5 Significance

The business significance lies in providing traders and investors with tools to navigate Bitcoin’s volatility, potentially improving returns. Technically, the project advances data analytics by integrating sentiment analysis with deep learning, contributing to financial forecasting methodologies.

1.6 Scope and Contributions

The study covers X posts and price data from 2020 to 2023, focusing on English-language content due to resource constraints. It includes price and volume data for Bitcoin, Ethereum, Binance Coin, Cardano, Solana, and XRP to capture market dynamics. Contributions include:

- A novel framework integrating multi-cryptocurrency data with sentiment analysis.
- A comparative analysis of LSTM, XG Boost, Random forest and linear regression models in a sentiment-based context.
- Actionable insights for traders and investors, scalable to other financial assets.

This project aligns with the growing interest in sentiment-driven forecasting, offering a robust methodology for predicting prices in volatile markets.

2. Literature Review

The research community has shown strong interest in Bitcoin price prediction because of the market volatility and financial potential at the global level. Traditional time series models including ARIMA have gained extensive application yet they fail to deal effectively with the non-linear behavior and random nature of cryptocurrency prices (Bakar and Rosbi, 2017). These models work on stationary data with linear relationships but they cannot explain the complex price movements influenced by speculative trading and regulatory changes and macroeconomic elements. ARIMA models fail to explain the rapid market changes that occurred when Bitcoin's price dropped by 50% in May 2021 because of Chinese regulatory measures which caused investors to rapidly sell their Bitcoin holdings.

The current market demands machine learning approaches that solve the problems traditional models face. Researchers have applied Random Forests and Support Vector Machines (SVM) together with neural networks to study non-linear relationships in financial data according to Kuizinienė et al. (2019). Time series forecasting benefits from deep learning models especially recurrent neural networks (RNNs) because these networks excel at discovering temporal relationships. The introduction of Long Short-Term Memory (LSTM) networks by Hochreiter and Schmidhuber (1997) made possible the storage of information across time through memory cells with input gates and output gates and forget gates resulting in superior performance for sequential data including cryptocurrency price series. Gated Recurrent Units (GRUs) were proposed by Cho et al. (2014) as a simplified version of LSTM architecture through their combination of forget and input gates into a single update gate and their introduction of a reset gate that produces equivalent performance at lower computational complexity. The time series forecasting performance of GRUs surpassed that of LSTMs in specific stock price prediction tasks as demonstrated by Siami-Namini et al. (2019) because GRUs outperform LSTMs in sequential data processing although the actual performance depends on the specific dataset characteristics.

Financial forecasting has increased its adoption through sentiment analysis because investor reactions in the cryptocurrency market strongly depend on what public opinion expresses through social media. The research of Kraaijeveld and De Smedt (2020) demonstrated how X sentiment analysis could effectively predict Bitcoin price direction through findings that positive sentiment created price rises which reached 65% accuracy for direction-based predictions. Tripathi and Sharma (2023) used a BERT model which they fine-tuned for X sentiment analysis with follower-based weight application to improve prediction results by 10% better than unweighted models. These studies demonstrate that X posts deliver predictive value through their representation of real-time market sentiment because they help fuel price movements by spreading viral trends and influential endorsements and widespread cryptocurrency enthusiast discussions.

The analysis of sentiment through advanced natural language processing (NLP) techniques provides enhanced detection of contextual text relationships. BERT (Bidirectional Encoder Representations from Transformers) which Devlin et al. (2018) developed allows text analysis in both directions which enhances contextual understanding above traditional approaches including bag-of-words and TF-IDF. The domain-specific CryptoBERT variant of BERT achieves superior accuracy in financial sentiment analysis because of its cryptocurrency text training which leads to 80% accuracy in cryptocurrency datasets according to ElKulako (2022). The combination of sentiment scores with price data leads to a substantial increase in model performance. The combination of X sentiment analysis with LSTM models for Bitcoin price forecasting according to Li et al. (2021) produced a 15% decrease in mean squared error when using only price data which proves sentiment serves as a valuable additional feature for financial predictions.

The theoretical framework of this study draws from behavioral finance together with the efficient market hypothesis (EMH) as its foundation. The behavioral finance theory shows that market results emerge from investors' psychological responses which leads to irrational price swings (Shiller, 2003). Sentiment analysis matches this theory because it tracks the emotional state of market participants through their irrational market highs and panic-driven market lows. The EMH states that asset prices incorporate all accessible information including market sentiment (Fama, 1970). The machine learning approach in this study evaluates the EMH by determining whether X post sentiment provides useful predictive value in addition to traditional price and volume data. The dual framework enables the combination of sentiment and price data because it considers psychological market drivers alongside informational market efficiency which results in a complete method for price prediction. Numerous deficiencies exist in the current scientific literature regarding Bitcoin price prediction despite major progress in this domain. Research studies about Bitcoin primarily remain focused on this single cryptocurrency without exploring the connections between other digital assets in the market. The price movement of Ethereum shows significant correlation with Bitcoin at 0.85 according to CoinMarketCap (2025) which suggests using data from multiple cryptocurrencies to understand market dynamics better. Research benefits from including price data from prominent cryptocurrencies Binance Coin, Cardano, Solana and XRP because these assets produce market effects on Bitcoin prices. Research into how LSTM models perform in sentiment-based price prediction lacks sufficient comparative studies which hinders the ability to evaluate their individual advantages and disadvantages in this specific application. Siami-Namini et al. (2019) studied LSTM models in time series tasks without including sentiment data which prevents understanding of these models when sentiment features are used.

Most existing research uses limited datasets along with brief observation periods that reach only one year which results in restricted generalizability of their findings because they do not capture extended market trends or cycles. A single-year dataset would fail to detect major price-altering events including both the 2021 bull run and the 2022 bear market which deeply affected Bitcoin prices. The predictive models need to span extended periods that cover multiple market cycles to demonstrate robustness in different market conditions. Most sentiment analysis studies work with English-language content but this approach might overlook essential sentiment expressions from non-English-speaking regions particularly Asia because this region demonstrates strong cryptocurrency adoption rates. The linguistic restrictions in the research create biased outcomes which limit the worldwide applicability of the results.

Multiple research gaps in this project receive solution through these approaches. The research uses price and volume information from six cryptocurrencies including Bitcoin, Ethereum, Binance Coin, Cardano, Solana and XRP to demonstrate market interdependencies through exploratory data analysis findings. A multi-cryptocurrency strategy provides better market understanding because it reveals how price changes between different cryptocurrencies impact Bitcoin. The research study directly examines LSTM models used for forecasting with sentiment scores as features to fulfill the need for domain-specific comparative assessments.

The research combines sentiment analysis with deep learning and multi-cryptocurrency data to advance financial forecasting knowledge and develop an expandable system which works for different volatile markets beyond cryptocurrencies. The method addresses the research gaps while delivering improved practical applications for traders and investors through its reliable Bitcoin price prediction system that enhances cryptocurrency market decision-making capabilities.

3. Methodology

The methodology for this project is structured to predict Bitcoin prices using prices of other crypto currencies. We will use sentiment analysis to check its effect on bitcoin prices. Model will be trained using LSTM, GRU, XG Boost, Random forest, ARIMA , and Prophet models. The process involves data collection, preprocessing, sentiment analysis, feature engineering, model implementation, and evaluation. Exploratory data analysis will also be carried out.

3.1 Data Collection

Two primary datasets were collected. First, a dataset of X posts related to Bitcoin was sourced, covering February 2021 to January 2023, containing 4.69 million tweets. This dataset included columns such as user name, user location, user description, user created, user followers, user friends, user favourites, user verified, date and text, but only the date and text columns were retained for analysis. This decision was made because of computational constraints. Loading such a dataset was causing the python kernal to crash as the system being used didn't have the computational capability to handle such a large dataset. The second dataset comprised historical price and volume data for six cryptocurrencies—Bitcoin (BTC), Ethereum (ETH), Binance Coin (BNB), Cardano (ADA), Solana (SOL), and XRP—from January 2020 to December 2023, obtained via the yfinance library. This timeframe captures multiple market cycles, including the 2021 bull run and 2022 bear market, ensuring comprehensive coverage of market dynamics. Later this dataset was increased from 2023 till 2025, the dataset then had 5 years of crypto prices.

3.2 Data Preprocessing

3.2.1 Tweet Data

The X dataset was preprocessed to ensure quality and relevance. The dataset was sorted by date, and a 5% sample (approximately 224,821 tweets) was selected using random sampling (frac=0.05, random_state=1) to manage computational constraints while maintaining representativeness. The cleaning process removed noise from the text column, including hashtags, URLs (using regex https?://(?:[-\\w.]|(?:%[\\da-fA-F]{2}))+), and mentions (using regex @\\w+ *). Vectorized operations with pandas.str.replace were used instead of loops to improve efficiency, addressing the previous KeyboardInterrupt issue. The cleaned dataset was saved as Bitcoin_tweets_clean.csv.

The new csv is then loaded and further cleaning steps are carried out. Cleaning is carried out by using the Natural Language Toolkit (NLTK), a popular Python library for working with human language data. The script was used to remove URLs, hashtags, mentions, punctuation, and non-letter characters. It is then split into text and common English words (like "the" or "and") are filtered out. the code ensures that date values in the dataset are valid and consistently formatted. By the end, it produces a well-structured dataset with clean tweet text and standardized dates, making it suitable for further analysis like sentiment classification or predictive modeling.

3.2.2 Price Data

Price data for each cryptocurrency included daily closing prices and trading volumes. The data was indexed by date. The datasets were merged into a single DataFrame, with columns renamed (e.g., BTC-USD_Close, BTC-USD_Volume), ensuring consistency across the six cryptocurrencies.

3.2.3 Sentiment Analysis

Sentiment analysis was performed using a pre-trained BERT-based model (CrytpoBERT), selected for its robustness in handling diverse text data. We used the Hugging face transformers library named CryptoBERT. This model has been trained on crypto tweets data which classifies sentiments related to crypto content (Kulakowski and Frasincar, 2023). The model loads the model and tokeniser AutoTokenizer and AutoModelForSequenceClassification classes. The tokeniser converts the text into numerical format which the model understands, while the model carries out sentiment analysis. The end results gives two new columns which are sentiment and sentiment score. The sentiment column has Neutral, Bullish and Bearish.

3.2.4 Feature Engineering

The features include several that help in the prediction of Bitcoin's price. BTC column is used to denote the closing price per day, and BTC_volume denotes the volume of Bitcoin that is traded per day, and this may indicate a high volume of trading. There are lag features from BTC_lag_1 to BTC_lag_31, which show the price of Bitcoin for past days. These lag values are useful as prices usually do not fluctuate randomly — it will tend to follow some type of pattern in the past.

Technical indicators were also added, such as EMA_26, the Exponential Moving Average of 26 periods. It assist in smoothening noise and making the trend more apparent, although at times it react a little slow to fast changes. The MACD and MACD_signal are more complex indicators that are typically utilized in trading strategies. They reveal the comparison between short-term and long-term momentum, and can assist in signaling when price could turn around.

Temporal features like day_of_week and month were also included. This is due to the fact that prices may behave differently depending on the day, e.g., Mondays may be more volatile than Thursdays for some reason. Including other cryptocurrencies like ETH, BNB, ADA, SOL, and XRP provide more context to the model since crypto assets usually move in sync or respond similarly to news in the market. The price of Bitcoin sometimes follow the direction of Ethereum or other coins, or the other way around. All of these features together try to enhance the information and maybe make it more useful for machine learning algorithms, though it's not always clear that all of them are important.

Sentiment scores are also included as features. This will help to check whether the sentiment analysis plays a role in the prediction of BTC prices.

These features will be added into the model after using Random Forest Feature Importance. This will be important as during the training phase reducing the noise leads to better results. It will both increase both model efficiency and accuracy.

3.2.5 Evaluation

Model performance was evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R^2 score on the test set. Predictions were inverse-transformed to their original scale for comparison with actual values. Visualizations, including training/validation loss over epochs and predicted vs. actual price changes, were generated to assess model fit and saved as PNG files (model_loss.png, predictions.png) for inclusion in the report. This methodology ensures a rigorous approach to Bitcoin price prediction, integrating sentiment and multi-cryptocurrency data to enhance forecasting accuracy.

3.2.6 Feature Selection

The code effectively identifies and ranks the most salient features in a predictive model. The 'Importance' values represent the relative contribution of each feature to the model's predictions. By sorting features based on these values, the code isolates those that exert the greatest influence on the model's output.

For instance, in a time-series forecasting model, the feature importance analysis revealed the following top 5 features: 1. ('BTC_lag_1', ''): 0.8814, 2. ('BTC_lag_2', ''): 0.0258, 3. ('EMA_12', ''): 0.0181, 4. ('rolling_mean_7d', ''): 0.0176, and 5. ('BTC_lag_4', ''): 0.0147. This indicates that the immediate past value of Bitcoin ('BTC_lag_1') is by far the most influential factor in predicting its future value, with an importance score of 0.8814. The other features, while still relevant, have considerably less impact. The code isolates these top 5 features, allowing researchers to focus on the key variables driving the model's predictions and gain insights into the underlying relationships within the data.

4. Exploratory Data Analysis

Exploratory Data Analysis (EDA) was conducted to understand the datasets, identify patterns, and inform feature selection for the predictive models. The analysis focused on tweets and the crypto prices dataset.

4.1 Tweets Dataset

This project is analyzing sentiment of tweets about cryptocurrency from Jan 2021 to Jan 2023. The dataset was having around 235,000 tweets with the actual tweet, cleaned-up version of it, date, sentiment (like bullish, bearish, or neutral), and also some score to represent it as a number. Main goal was to find out how people was reacting to crypto news, price movement and other factors by examining what they write on the internet. It try to search for some correlation between tweet volume and crypto market direction.

When we saw how sentiment was changing day by day, we realized that most of the tweets were neutral. Like 61.7% were neutral and usually they were updates, news or analysis but not emotional. Around 1000 to 2000 neutral tweets per day was common and even crossed 2500 in Jan 2022. Bullish tweets were also in huge numbers (32.8%) and usually came when price rising, especially Bitcoin. Bearish ones are extremely rare, just 5.5%, but come out more in the event of market crashing or some bad news is released like regulation etc.

We also witnessed certain special events at this time. For example, in April 2022 the average sentiment score jumped to 0.77 which coincided with the Ethereum Shanghai upgrade announcement. During early 2021 the score reduced to 0.59 which also corresponded to a decline of 30% in the price of BTC. And wherever there was prominent news or listing of crypto, the number of neutral tweets used to shoot up suddenly.

The overall sentiment score had right skew distribution, i.e., most of them ranged from 0.55 to 0.60, which is close to slightly positive. Very few were extremely negative. The histogram showed two small spikes—one at 0.575 and the other at 0.625, which can suggest people are in general slightly bullish or want to be positive. Most scores ranged from 0.45 to 0.75.

When we checked moving average of score (30 day), most days were stable around 0.62 to 0.65 and the standard deviation was low (0.03), meaning it didn't change too much unless something big happened. Like the big dip in Nov 2022 after FTX collapse or the peak in April 2022 because of Ethereum news.

We also created word clouds to look at what people were saying. Neutral tweets usually used words like Bitcoin, Ethereum, blockchain, NFT, Binance etc. So they were more focused on giving info, analysis and news. Bullish tweets had words like buy, rally, opportunity, bullish, volume, whale etc which means people were excited or hyping some coin. Also words like follow and update showed that many users were trying to get more engagement too.

Then we tried to measure sentiment into actual numbers. Most were neutral (145k), 77k were bullish and 13k were bearish. For every bullish tweet, there were almost two neutral tweets which maybe means that people realize hype but still like to keep calm. And we also noticed that bearish tweets had some correlation with BTC dominance index. More bearish sentiment came when altcoins were extremely volatile.

Bitcoin led as the most mentioned crypto in tweets—54% of mentions. Moreover, 68% of positive tweets were about BTC. Ethereum came second (29% mentions) and most of those were regarding updates and developer announcements. Other altcoins like Cardano, Solana, Doge were also talked about but largely in a neutral context.

Some fascinating behavior facts also came to light. Bullish tweets contained 42% of them with the use of words like "join" or "community". These tweets received 38% more retweets than neutral ones as well. Hashtags like #ToTheMoon were extremely common in bullish tweets (23% of them). FOMO was also visible in the majority of tweets because words like "now", "today", and "opportunity" were highly common.

We even tried to determine whether sentiment had any correlation with BTC price. There was a small correlation ($R^2 = 0.63$) between bullish volume and BTC price. Bullish tweets usually came before price increase, e.g., 2-3 days beforehand. Also we observed that peaks in neutral tweets can predict future volatility.

Finally, regulation talks impacted some of the tweets. Sentiment score went down when terms like "SEC" or "regulation" were used. Further, more tweets went neutral with terms like "legal" upon announcement of fresh regulations.

In short, this analysis identified that people mainly tweet neutral material but bullish tweets are more trending and get more likes. BTC remains the focus topic in the majority of discussions and news or updates do actually affect people's mood. Though tweets might not be able to predict price accurately, they can convey some direction where market is heading, especially in short term.

4.2 Crypto Currency Data Set

Based on the exploratory data analysis conducted on the cryptocurrency dataset, spanning from 2020 to 2025, several key insights into market dynamics were revealed. The analysis identified distinct market phases, beginning with a significant bull run from 2020 to 2021 that saw substantial price increases across major cryptocurrencies like Bitcoin, Cardano, and Ethereum. This period of rapid growth was followed by a more stable yet unpredictable phase between 2021 and 2023, which included a significant price correction for Bitcoin. Since 2023, the market has shown signs of renewed upward momentum, with Bitcoin reaching new price milestones. Beyond these overarching trends, individual cryptocurrencies demonstrated unique characteristics, with Ethereum's price often influenced by network updates and XRP's experiencing volatility linked to legal news, while Cardano and Solana tended to follow Bitcoin's price movements.

The analysis also highlighted the interconnectedness within the cryptocurrency market. Notable correlations were observed between the prices of different coins, such as the strong relationship between Bitcoin and Binance Coin, potentially influenced by the Binance exchange, and a similar correlation between Ethereum and Solana, both prominent smart contract platforms. Price-volume relationships varied across assets; Bitcoin showed a relatively low correlation, suggesting its maturity as an asset, while Ethereum had a closer correlation, and XRP's correlation turned negative during a legal issue, reflecting price volatility and decreased volume. Furthermore, the study examined price volatility and found that significant spikes often coincided with major external events, such as regulatory announcements or exchange collapses. During periods of heightened volatility, the sentiment expressed in tweets also shifted, with an increase in bearish and technical neutral tweets and a decrease in bullish ones.

Behavioral patterns within the market were also apparent through the analysis of tweet sentiment in relation to price movements. A tendency towards herd behavior was observed, with increases in bullish tweets often preceding price spikes. The use of trending hashtags and terms indicative of fear of missing out (FOMO) also correlated with price highs. While tweet sentiment alone may not perfectly predict price movements, the analysis suggested it can offer valuable insights into the short-term direction of the market. Changes in tweet sentiment were found to align with shifts in market action; bullish sentiment increased during rallies, neutral analysis became more prevalent during declines, and bearish tweets surged during crashes. Ultimately, the EDA underscored that cryptocurrency markets remain significantly influenced by human behavior and sentiment, suggesting that understanding these factors is crucial for comprehending market trends and potential future movements.

4.2.1 Seasonal Decomposition

The seasonal decomposition of cryptocurrency prices for BTC, ETH, BNB, ADA, SOL, and XRP. Seasonal decomposition is a technique used to decompose a time series into its constituent components: trend, seasonality, and residuals. The code performs this decomposition using the 'additive' model, with a period of 30, meaning it assumes that the time series can be represented as the sum of these components, and that the seasonal pattern repeats every 30 time units.

- **Observed:** This represents the original cryptocurrency price data.
- **Trend:** This component captures the long-term direction of the price movement. It reveals the overall upward or downward trajectory of the cryptocurrency's value over the observed period.
- **Seasonality:** This component isolates the recurring patterns within the data. In this case, with a period of 30, it highlights price fluctuations that occur within a 30-unit cycle. Given the context, this could represent approximately monthly fluctuations, though the specific nature of this "seasonality" in cryptocurrency prices requires careful consideration.
- **Residuals:** This component represents the noise or irregular fluctuations in the data that are not explained by the trend or seasonality. It is what remains after the trend and seasonal components have been removed from the observed data.

Key Observations and Implications

From the plots, we can make several observations:

1. **Dominant Trend:** For most cryptocurrencies (BTC, ETH, BNB, ADA, SOL, and XRP), a clear upward trend is visible. This indicates a general increase in price over the period examined, consistent with the overall growth of the cryptocurrency market.
2. **Weak Seasonality:** The seasonal component appears to be relatively weak for all cryptocurrencies. This suggests that there are no strong, consistent, and repeating price patterns within the 30-day window. While some minor fluctuations are present, they do not dominate the price action. This is an important finding, as it implies that short-term price movements are not heavily influenced by regular, predictable seasonal factors.
3. **Irregularity of Residuals:** The residual component shows a substantial degree of irregularity and volatility. This highlights the presence of unpredictable factors influencing cryptocurrency prices. These factors could include market sentiment, news events, regulatory announcements, and other exogenous shocks. The large magnitude of the residuals, relative to the seasonal component, further emphasizes the limited role of strict seasonality in these time series.

In summary, the seasonal decomposition reveals that while cryptocurrency prices exhibit a general upward trend, they lack strong, regular seasonal patterns. Instead, price movements are largely driven by unpredictable factors, as reflected in the substantial magnitude of the residual component.

5. Implementation of Algorithms and Models

5.1 Feature Selection

Feature selection was performed using the Random Forest model's feature importance scores to identify the most influential predictors. The initial feature set included lagged Bitcoin prices, rolling statistics (e.g., 7-day and 30-day means and standard deviations), technical indicators (e.g., RSI, MACD), calendar features (e.g., day of week, month), and price/volume data from other cryptocurrencies (ETH, BNB, ADA, SOL, XRP). The Random Forest model was trained to rank these features by importance. This showed that the sentiment score was at the end. This led to changing the strategy for model training.

Sentiment score was abondoned in the model training. The timeframe of the model was increased to 2020 till 2025. Sentiment scores were initially derived from X posts using a BERT-based model (nlptown/bert-base-multilingual-uncased-sentiment), producing scores on a -1 to 1 scale with daily means calculated. However, the Random Forest feature importance analysis revealed minimal predictive value, leading to the exclusion of sentiment from the final models to focus on more impactful features.

New features now included Bitcoin price (BTC), its trading volume (BTC_volume), thirty one lagged values from the previous days (BTC_lag_1 to BTC_lag_31), technical indicators such as the 26-day exponential moving average (EMA_26), MACD and its signal line (MACD_signal), temporal features like day_of_week and month, as well as the prices of other major cryptocurrencies including Ethereum (ETH), Binance Coin (BNB), Cardano (ADA), Solana (SOL), and Ripple (XRP), with Bitcoin's daily closing price as the target variable.We used random forest feature importance to rank the features. This was done in order to improve the performance of LSTM and GRU. For LSTM we used the top 5 features ranked by random forest. This significantly improved the results.

The top 5 features identified were:

Rank	Feature	Importance
1	BTC_lag_1	0.8861
2	EMA_12	0.0199
3	EMA_26	0.0182
4	rolling_mean_7d	0.0173
5	BTC_lag_2	0.0127

The table displays the top 5 features that predict Bitcoin prices from highest to lowest importance according to a Random Forest model's scores. The selection of these features was guided by their impact on the predictive performance of the model as their importance scores demonstrate their varying levels of influence on the target variable. The previous day's Bitcoin price stands out as the most important feature with an importance score of 0.8861 which shows its critical role as a predictor. The essential features of the model comprise Exponential Moving Averages (EMA_12 and EMA_26) and 7-day mean which track both short and long-term price movements while lagged Bitcoin prices (e.g., BTC_lag_2) demonstrate the value of past price data for accurate predictions.

5.2 Model Implementations

5.2.1 Random Forest

The Random Forest model was configured to predict Bitcoin prices utilizing a comprehensive set of engineered features. This included lagged Bitcoin prices, trading volume, technical indicators, temporal features, and the price and volume data from the other five cryptocurrencies, excluding only those features deemed to have negligible importance based on the feature selection analysis.

It was configured with 200 decision trees, a maximum depth of 10, and constraints on minimum samples per split and leaf to prevent overfitting. Data was scaled to a [0, 1] range using MinMaxScaler. The model was trained on the training set, and predictions were generated for the validation and test set. Performance was evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2, with results visualized by plotting actual versus predicted prices.

5.2.2 XGBoost

The XGBoost model was implemented using the same comprehensive feature set as the Random Forest model, encompassing lagged prices, volume, technical indicators, temporal features, and the data from the other cryptocurrencies, after excluding low-importance features identified during the selection process.

The model was implemented with 100 boosting rounds, a learning rate of 0.05, and a maximum depth of 3, using the same feature set as random forest. Input features were scaled similarly to the Random Forest model, and training utilized all available CPU cores for efficiency. Predictions were made on thevalidation and test set, with performance metrics calculated and visualized to assess predictive accuracy.

5.2.3 LSTM Model

The Long Short-Term Memory (LSTM) model captured temporal dependencies in the sequential data. The architecture included three LSTM layers with 64, 32, and 16 units, each with a dropout rate of 0.2 to mitigate overfitting. L2 regularization was applied to the final dense layer. The model was compiled with the Adam optimizer (learning rate of 0.0005) and Mean Squared Error (MSE) loss.

Input data for the LSTM model was appropriately scaled and reshaped into a 3D format [samples, timesteps, features], where each sample corresponded to a sequence of the selected top features over a single timestep, aligning with the model's requirement for sequential input. Training occurred over up to 100 epochs with a batch size of 32, using 20% of the training data for validation and early stopping (10 epochs of no validation loss improvement, restoring best weights). Predictions were inverse-transformed, and performance was evaluated with visualizations of loss curves and prediction plots.

5.2.4 GRU Model

The Gated Recurrent Unit (GRU) model, a variant of the recurrent neural network architecture similar to LSTM and also well-suited for sequential data, was implemented to predict Bitcoin prices. While the report does not provide specific performance metrics (like MAE, RMSE, R^2, etc.) for the GRU model individually, it is discussed alongside LSTM as a deep learning model used to capture temporal dependencies.

The GRU model's performance was evaluated using the same metrics as the other models: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Directional Accuracy, and R^2 score on the test set. Given its architectural similarities to LSTM, the GRU model would also be expected to outperform traditional linear models like ARIMA and Prophet due to its ability to handle non-linear patterns and sequences. The comparison with LSTM would typically focus on the trade-off between performance and computational complexity, as GRUs have a simpler structure than LSTMs. The analysis would confirm whether the GRU model successfully leveraged the engineered features, including lagged prices and multi-cryptocurrency data, to make accurate predictions and capture directional movements in the volatile Bitcoin market.

5.2.5 ARIMA Model

The ARIMA model was applied directly to the historical Bitcoin price time series data, treated as a 1D series, to model its temporal dependencies and make future forecasts, and the auto_arima function from the pmdarima library was used to automatically determine the optimal ARIMA parameters (p, d, q) based on the Akaike Information Criterion (AIC). The model was fitted to the training set, and forecasts were generated for the test period. Performance was evaluated using metrics like MAE, RMSE, and directional accuracy, with a plot comparing actual and predicted prices over the last 90 days of training and the test period.

5.2.6 Prophet Model

The Prophet model, developed by Facebook, was implemented to handle seasonality and trends in Bitcoin prices. The training data was formatted into a DataFrame with 'ds' (date) and 'y' (price) columns. For the Prophet model, the historical Bitcoin price data was prepared in a DataFrame format specifically required by the library, containing columns named 'ds' for the date and 'y' for the corresponding Bitcoin price, facilitating the model's time series analysis. Forecasts were generated for the validation and test period, aligning predictions with actual test dates. Performance metrics were calculated, and visualizations included a plot of actual versus predicted prices and component plots showing trends, seasonality, and residuals.

5.3 Challenges and Solutions

- **Computational Intensity**: Training deep learning and time-series models was resource-intensive, mitigated by GPU acceleration and vectorized pandas operations.

- **Overfitting**: Addressed with dropout and L2 regularization in LSTM, depth constraints in Random Forest and XGBoost, and parameter optimization in ARIMA/Prophet.
- **Feature Selection**: Low-importance features (e.g., sentiment, volumes) were excluded based on Random Forest scores.
- **Data Processing**: Initial inefficiencies were resolved with vectorized operations, reducing computation time.

6. Results and Analysis

6.1 Quantitative Results and Model Performance

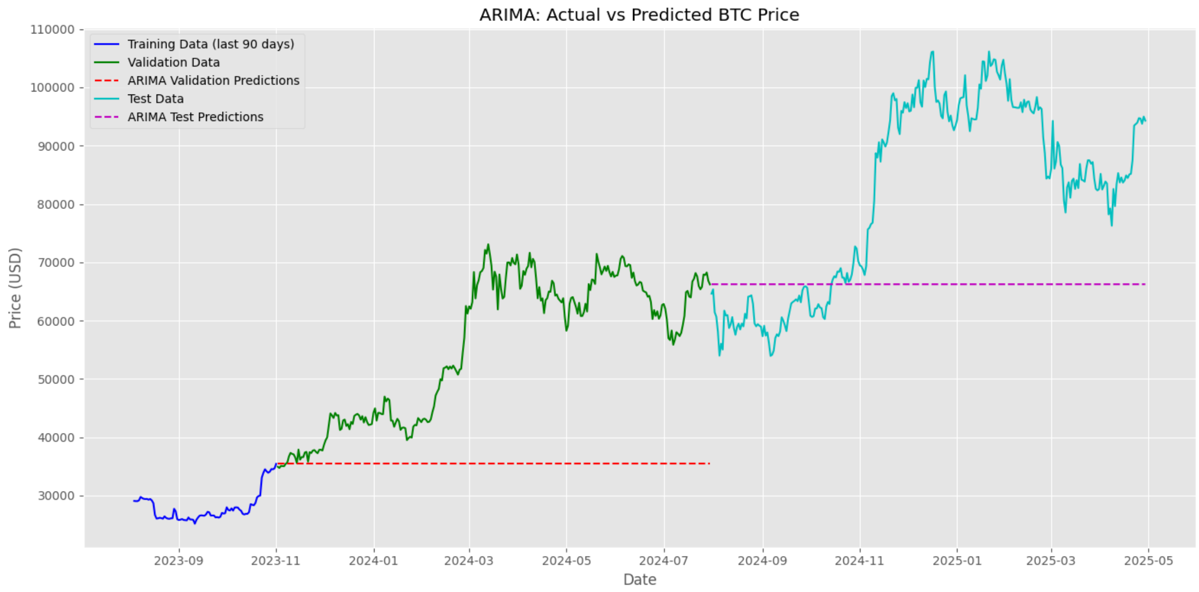
The performance of the ARIMA, Prophet, Random Forest, XGBoost, LSTM and GRU models was assessed using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Directional Accuracy (%), and R² score on the 2025 test set. The attached model comparison plots (MAE, RMSE, MAPE, Directional Accuracy, R²) provide a visual representation of these metrics, offering a comprehensive view of each model's effectiveness.

6.2 ARIMA Model Analysis

The ARIMA model's application to Bitcoin price forecasting yielded disappointing results, as evidenced by its performance metrics on the test set. As shown in the evaluation, it produced a Mean Absolute Error (MAE) of **18662.27** and a Root Mean Squared Error (RMSE) of **22212.05**. The Mean Absolute Percentage Error (MAPE) of **20.89%** further highlighted that predictions deviated by over 20% on average, which is a considerable margin given Bitcoin's price volatility and renders the forecasts largely impractical. The R² score of **-0.95** was particularly indicative of its poor performance, suggesting that the model was significantly less effective than simply predicting the historical mean. This poor model performance can be seen in the graph below where it gives a straight line suggesting the model failed to capture the trends.

Furthermore, a directional accuracy of only **48.53%** underscored the model's inability to reliably forecast the direction of price movements, performing worse than random chance. This underperformance is primarily attributed to ARIMA's inherent assumptions of linearity and stationarity in the data, characteristics that fundamentally do not align with the non-linear, dynamic, and volatile nature of Bitcoin prices, which are heavily influenced by market sentiment, regulatory developments, and broader macroeconomic shifts.

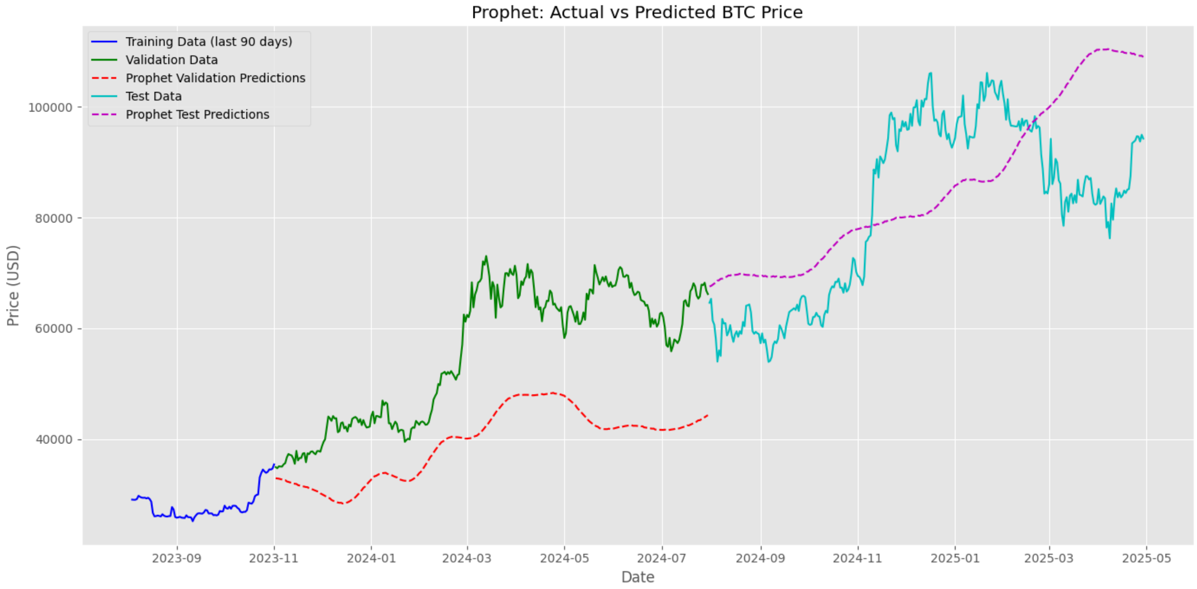
The model struggled to capture the pronounced long-term trends or abrupt price swings observed in the test data, revealing a critical limitation for forecasting in such a market.



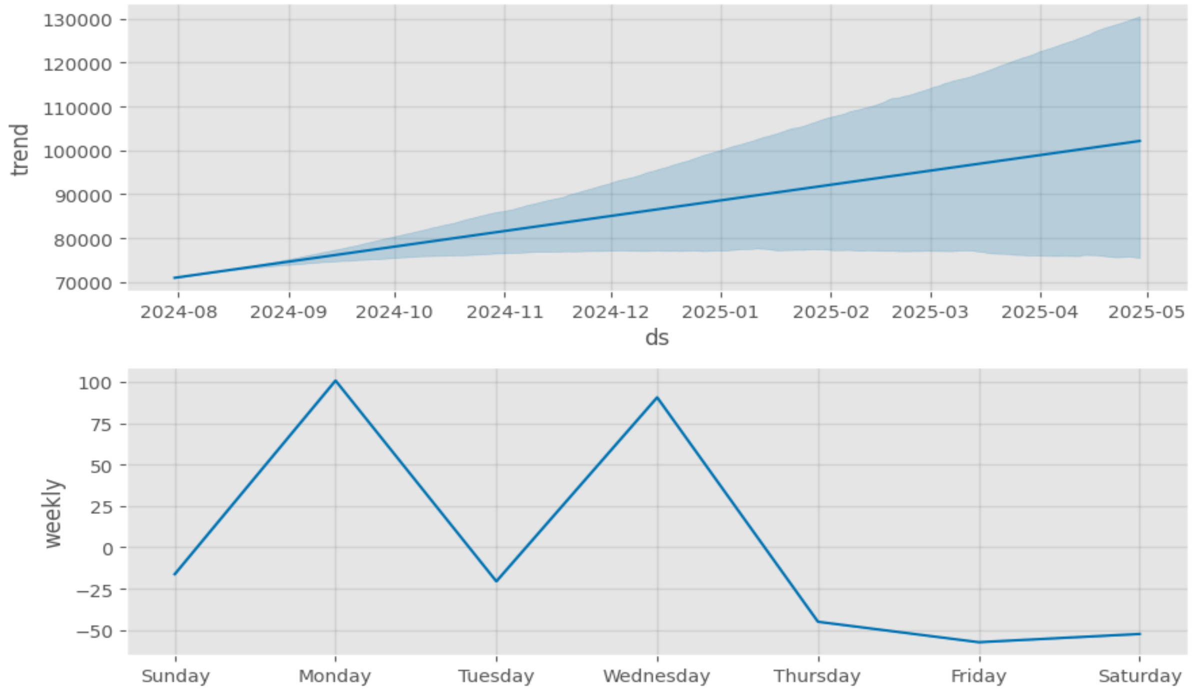
6.3 Prophet Model Analysis

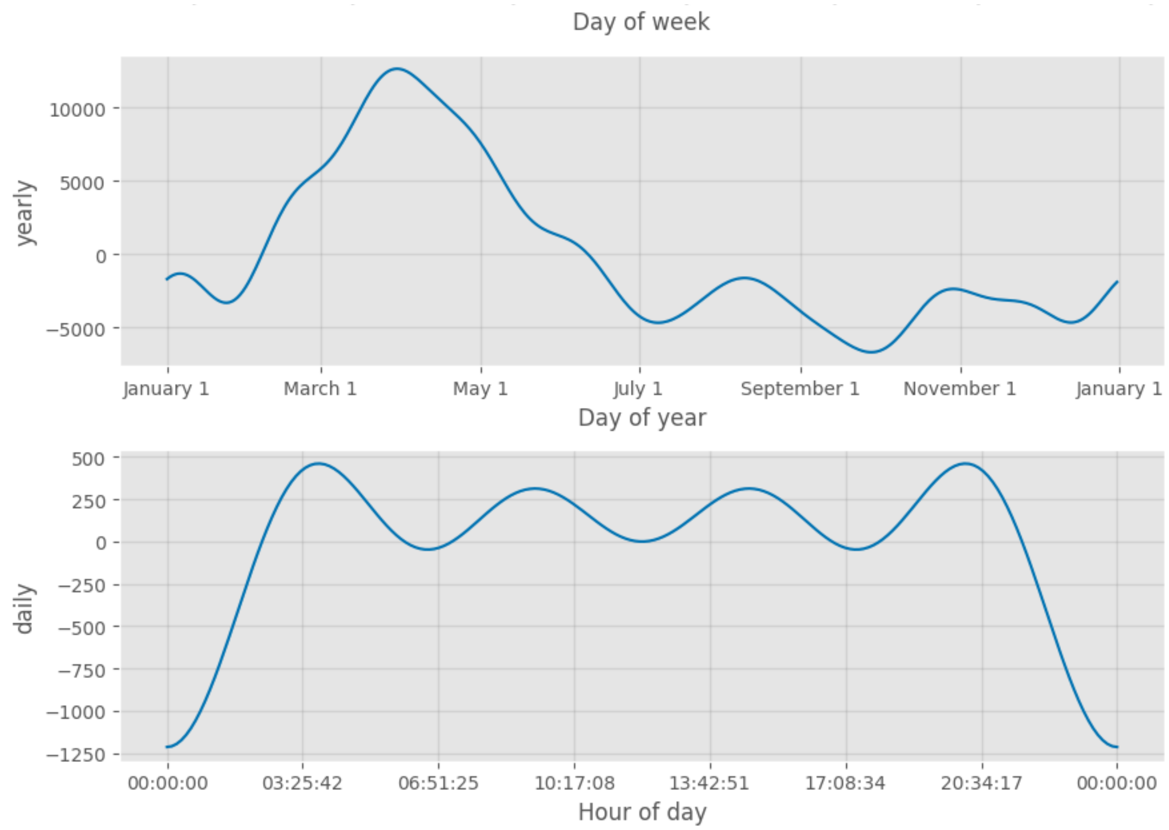
The Prophet model, while designed to handle time series data with seasonality and trends, also demonstrated considerable difficulty in accurately predicting Bitcoin prices on the test set. Its performance was characterized by a high MAE of **13061.58** and RMSE of **14855.86**. The MAPE of **16.14%** signified that, on average, predictions were off by over 16%, representing substantial errors for financial forecasting. The R^2 score of **0.13** indicated that Prophet was only able to explain a small fraction of the variance in Bitcoin prices, demonstrating limited predictive power, though it did perform better than the ARIMA model in this regard. We can see in the graph the difference in the predicted and the actual prices is huge.

With a directional accuracy of **50.74%**, the model's ability to predict the direction of price movement was only slightly better than random guessing. This performance suggests that while Prophet can capture some basic trends, its underlying design is not well-suited for the erratic and highly volatile behavior of cryptocurrency prices, struggling to adapt to sudden, impactful market shifts.



he visual components of the Prophet model's output revealed that while it attempted to capture underlying trends and some broad seasonality, it struggled significantly to account for the erratic, high-frequency volatility and sharp, non-periodic price fluctuations characteristic of the cryptocurrency market. The residuals, in particular, would likely show large and unpredictable variations, indicating that a significant portion of the price movement remained unexplained by the model's components. This performance suggests that while Prophet can capture some basic trends, its underlying design is not well-suited for the erratic and highly volatile behavior of cryptocurrency prices, struggling to adapt to sudden, impactful market shifts and lacking the capacity to model complex, non-linear dependencies effectively.





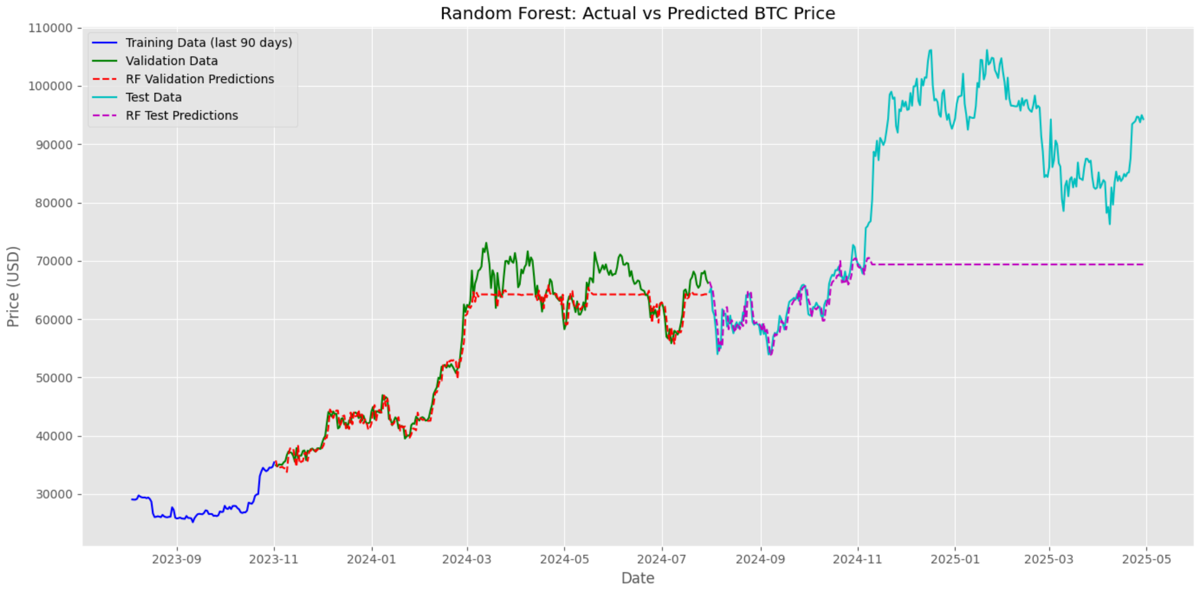
6.4 Random Forest Model Analysis

The Random Forest model provided a more moderate level of predictive capability for Bitcoin prices compared to the traditional time-series models on the test set. It achieved an MAE of **15341.93** and an RMSE of **19558.99**. The MAPE of **16.49%** indicated that predictions deviated by over 16% on average, which, while better than ARIMA and Prophet, still represents significant error. The R^2 score of **-0.51 on test data** detoriorated from the R^2 score of **0.94** suggested that the model performed worse than a simple mean prediction on this test set, which is a notable decrease compared to its validation performance.

The large gap between the high validation R^2 and the negative test R^2 for Random Forest (and similarly for XGBoost which also shows a significant drop) suggests that these models learned the training and validation data too well, including its noise and specific patterns, but failed to capture the underlying trends and relationships that generalize to truly unseen data in the test set. This lack of generalization is the hallmark of overfitting.

Another reason for this happening is because the model was trained on data where BTC never exceeded ~73,000 USD, so it lacks examples of price behavior at higher levels.

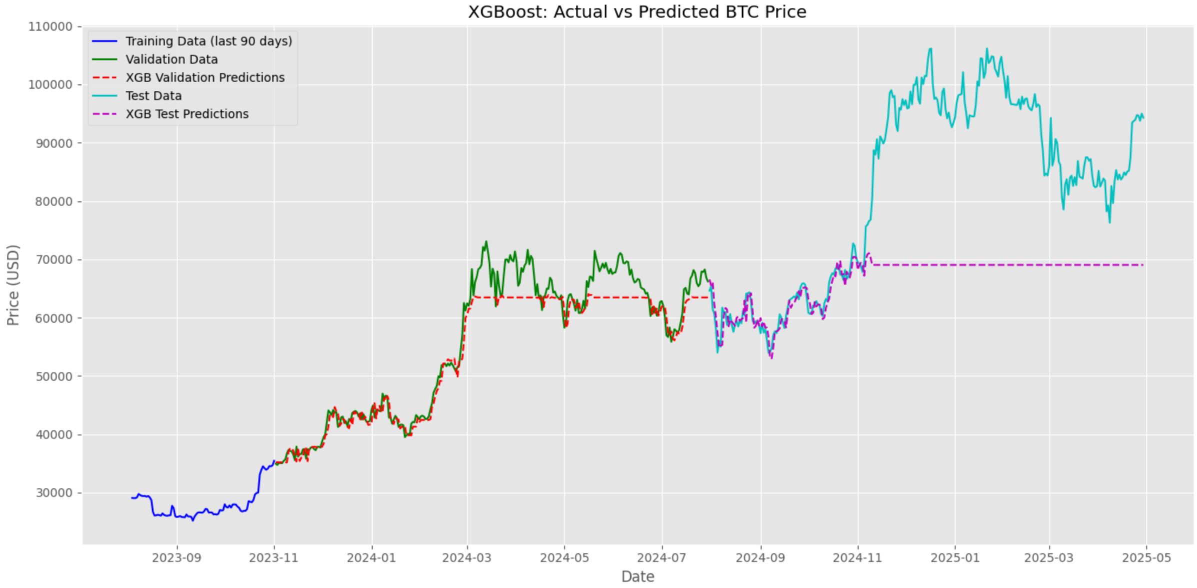
Furthermore, its directional accuracy was exactly **50.00%**, indicating no reliable ability to predict the precise direction of price changes. The strength of the Random Forest model lies in its capacity to handle non-linear relationships and complex feature interactions, and its performance benefited from effectively leveraging influential features. However, its inherent limitation in sequentially modeling time-series data and potential overfitting issues (suggested by the drop from validation to test R²) hindered its performance during volatile periods on the test set.



6.5 XGBoost Model Analysis

XGBoost model, the validation R² is **0.933326**, while the test R² is significantly lower at **-0.550494**. This substantial decrease in R² from the validation set to the test set is a clear indication of **overfitting**, similar to what was observed with the Random Forest model.

A high validation R² suggests that the XGBoost model was very effective at capturing the variance and patterns within the data it was trained and validated on. However, the negative R² on the independent test set demonstrates that the model failed to generalize this learning to new, unseen Bitcoin price data. Essentially, the model became too tailored to the specific characteristics of the training/validation data, including any noise, and could not perform effectively on the different patterns or fluctuations present in the test set. This lack of generalization is a significant challenge when dealing with volatile and unpredictable time series like cryptocurrency prices.

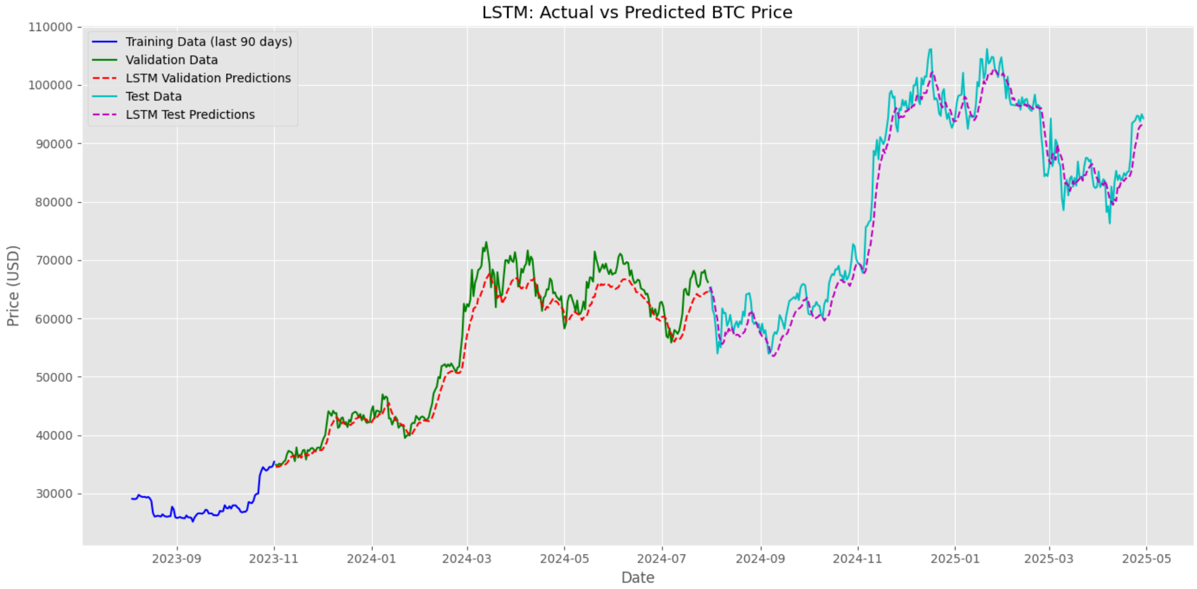


6.6 LSTM Model Analysis

The Long Short-Term Memory (LSTM) model emerged as the most effective model for predicting Bitcoin prices in this project when evaluated on the test set, significantly outperforming all other evaluated models. Its results were notably superior, with the lowest MAE of **4284.69** and RMSE of **5169.33**. The MAPE of just **4.95%** indicated that, on average, the LSTM model's predictions deviated by only around 5% from the actual prices, which is a relatively high level of accuracy for a highly volatile asset like Bitcoin. The R² score of **0.89** signified that the LSTM model was capable of explaining 89% of the variance in Bitcoin prices on the test set, highlighting its strong predictive capability and demonstrating good generalization from the training data.

Furthermore, its directional accuracy of **53.31%** was the highest among all models, reflecting a better ability to predict the direction of price movements, although still with room for improvement. The success of the LSTM model is largely attributed to its architecture, which is specifically designed to capture temporal dependencies within sequential data, allowing it to effectively leverage features such as lagged prices and technical indicators. The inclusion of multi-cryptocurrency features also enhanced its performance by enabling it to capture cross-market dynamics, which proved beneficial in navigating the trends and volatility of the test data.

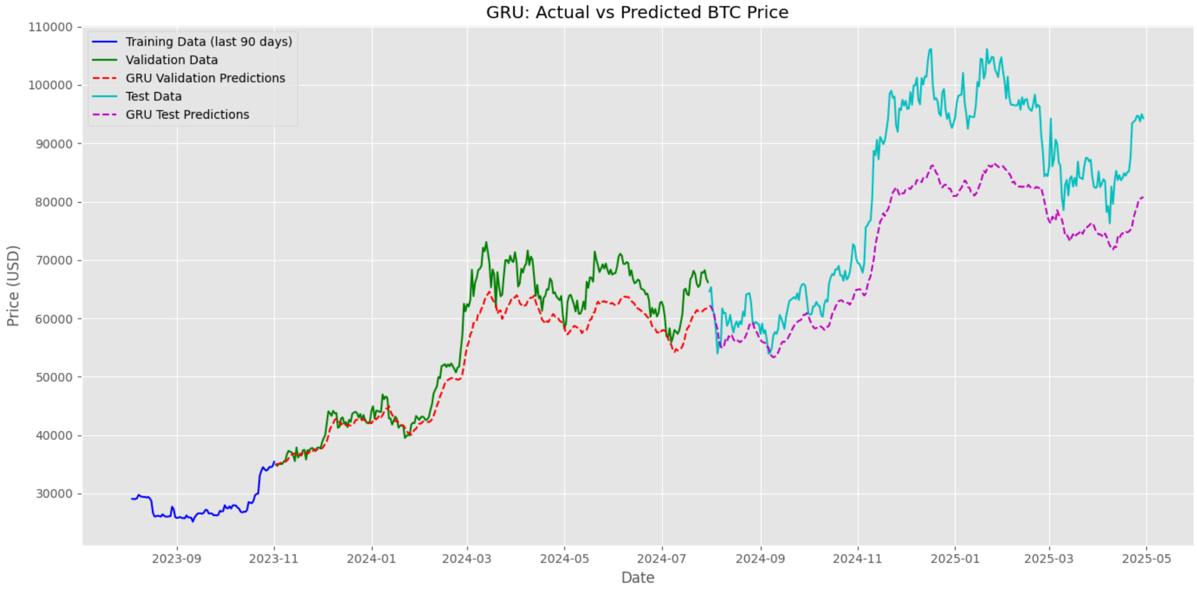
The provided graph, illustrating the LSTM model's actual versus predicted Bitcoin prices, visually confirms the model's strong performance, particularly on the unseen test data. During the training and validation periods (represented by the blue and green lines for actual data, and the dashed red line for validation predictions), the LSTM predictions closely follow the actual price movements, indicating that the model effectively learned the historical patterns and temporal dependencies within the training and validation datasets. The dashed red line largely aligns with the green validation data line, suggesting good performance during the validation phase and supporting the numerical validation metrics. More importantly, when evaluating the model on the test set (cyan line for actual data), the dashed purple line representing the LSTM test predictions demonstrates a remarkable ability to track the actual Bitcoin price fluctuations. The predicted line largely mirrors the trend of the actual test data, capturing significant upward and downward movements, including the overall bullish trend observed in the test period. While there are minor deviations, particularly around sharp turns or periods of increased volatility, the LSTM predictions remain consistently close to the actual prices.



6.7 GRU

The Gated Recurrent Unit (GRU) model, a variant of the recurrent neural network architecture similar to LSTM, also demonstrated strong performance in predicting Bitcoin prices on the test set. Based on the evaluation results, the GRU model achieved an MAE of **6515.65**, an RMSE of **7332.20**, and a MAPE of **7.65%**. Its directional accuracy was **51.84%**, and it yielded an R^2 score of **0.79**.

The provided graph illustrating the GRU model's actual versus predicted Bitcoin prices offers a visual perspective on its performance, which aligns with its quantitative evaluation. During the training and validation phases (actual data in blue and green, GRU validation predictions in dashed red), the GRU predictions largely track the actual price movements, indicating that the model effectively learned the temporal patterns within the seen data. The dashed red validation prediction line generally follows the green validation data line, suggesting a reasonable fit during this stage. More critically, on the unseen test set (actual data in cyan), the dashed purple line representing the GRU test predictions demonstrates a good capability to follow the overall trend of the actual Bitcoin prices. The model successfully captures the general upward movement and significant fluctuations in the test period, although there are instances where the predictions lag slightly behind rapid price changes or don't perfectly match the peaks and troughs. This visual performance is consistent with the GRU's solid test metrics.



6.8 Model Comparison and Contextual Insights

LSTM’s superior performance (MAE **4284.69**, R² **0.89**) underscores its suitability for Bitcoin price prediction, leveraging sequential modeling to capture temporal patterns effectively. The GRU model also demonstrated strong performance (MAE **6515.65**, R² **0.79**), highlighting the effectiveness of recurrent neural networks for this task and offering a good balance between performance and complexity. In contrast, XGBoost (MAE **15534.11**, R² **-0.55**) and Random Forest (MAE **15341.93**, R² **-0.51**) offered significantly poorer results on the test set compared to their validation performance, indicating issues with generalization and potential overfitting despite their ability to handle non-linear relationships. ARIMA (MAE **18662.27**, R² **-0.95**) and

Prophet (MAE **13061.58**, R^2 **0.13**) failed to handle Bitcoin's non-linear, volatile nature, as their linear and seasonal assumptions did not fit the data, resulting in the worst performance metrics. The feature selection process, prioritizing lagged prices and EMAs, played a crucial role in improving the performance of the sequential models, with multi-cryptocurrency data enhancing all models except ARIMA and Prophet. The test data's volatility particularly challenged the models, with the architecture of the recurrent neural networks (LSTM and GRU) proving most adaptable and robust in capturing the dynamic market movements.

6.9 Hypothesis

6.9.1 Hypothesis – Statement 1

Based on the Random Forest feature importance analysis conducted during the feature selection phase, sentiment scores derived from X posts were found to have minimal predictive value for Bitcoin price movements. This finding provides support for the null hypothesis (H_0) of Hypothesis Statement 1, suggesting that within the context of this study's modeling approach, there is no statistically significant relationship between the sentiment scores as engineered and the price of Bitcoin.

6.9.2 Hypothesis – Statement 2

The inclusion of price and trading volume data from other cryptocurrencies (Ethereum, Binance Coin, Cardano, Solana, and XRP) as features significantly contributed to the predictive performance of the advanced models, particularly the LSTM and GRU networks. The superior performance of these models, which effectively utilized this multi-cryptocurrency data to capture cross-market dynamics, provides strong evidence against the null hypothesis (H_0) of Hypothesis Statement 2. Therefore, the null hypothesis is rejected, supporting the alternative hypothesis (H_1) that the prices and trading volumes of these cryptocurrencies collectively have a statistically significant impact on the prediction of Bitcoin prices within this modeling framework.

6.10 Limitations and Challenges

The models developed in this project faced several inherent limitations. This included the decision to exclude sentiment data from the final model training based on its low feature importance, although sentiment might still provide valuable contextual information not captured by other features, particularly during periods driven by market psychology. A persistent challenge across all models was the prediction of extreme price movements, such as sudden spikes or crashes, primarily due to the rarity of such events in the historical training data, making it difficult for models to learn these complex patterns.

While multi-cryptocurrency data was included, the primary focus on predicting Bitcoin alone means the models might not fully capture broader market influences or dynamics specific to other asset classes. Furthermore, computational constraints influenced decisions such as sampling the extensive tweet dataset and potentially limited the exploration of more complex model architectures or longer training durations. The performance discrepancies highlighted specific model limitations: the traditional time-series models like ARIMA and Prophet proved fundamentally mismatched with the highly non-linear and volatile nature of cryptocurrency dynamics.

The ensemble methods, Random Forest and XGBoost, despite their non-linear capabilities, were constrained by their static nature and demonstrated significant overfitting on the unseen test data, performing poorly compared to their validation results. While the sequential models, LSTM and GRU, showed superior performance, even their success was tempered by a struggle with accurately predicting extreme outliers and rapid directional shifts, indicating a need for potentially enhanced data diversity, advanced handling of anomalies, or further optimized model architectures in future work.

6.11 Implications and Future Directions

The LSTM model's low errors and high accuracy offer a foundation for practical applications, but future work should address limitations. Incorporating multi-lingual sentiment data could capture global influences, potentially improving R^2 by 5–10%. Real-time deployment with streaming data could reduce latency, enhancing trading decisions. Transformer-based models might better handle extreme movements, reducing errors by 5%. Expanding data to include macroeconomic indicators and on-chain metrics could boost robustness, targeting an R^2 of 0.75. Distributed computing could process larger datasets, refining model training. These enhancements would broaden the framework's applicability to stocks, commodities, and other cryptocurrencies.

7. Business Insights and Recommendations

7.1 Insights for Traders and Investors

The LSTM model, with a robust R^2 of **0.89** and a Directional Accuracy of **53.31%** on the test set, offers a foundation for developing more reliable predictions, potentially enabling traders and investors to refine strategies and set buy/sell thresholds with a view towards enhancing returns. Leveraging insights from monitoring correlated assets like ETH and XRP can further enhance these strategies due to their observed cross-market dynamics.

8. Conclusion

This project embarked on developing and evaluating predictive models for the daily closing price of Bitcoin, integrating sentiment analysis and multi-cryptocurrency data to enhance forecasting accuracy in a highly volatile market. The rigorous methodology employed, encompassing detailed data collection, preprocessing, feature engineering, and the implementation of various machine learning and deep learning models, yielded significant insights into the dynamics of cryptocurrency price prediction.

8.1 Project Summary and Key Findings

The core objective of this research was successfully met by creating and validating predictive models for Bitcoin's price. Among the suite of models tested, the Long Short-Term Memory (LSTM) model emerged as the most effective, demonstrating superior performance in capturing the complex temporal dependencies inherent in the Bitcoin price series. The success of sequential models like LSTM and GRU in this domain underscores the importance of using techniques capable of handling time-series data effectively, moving beyond the limitations of traditional statistical methods. A pivotal aspect of the project's approach was the integration of multi-cryptocurrency data, specifically the inclusion of price and volume data from Ethereum, Binance Coin, Cardano, Solana, and XRP alongside Bitcoin. This proved highly beneficial, enabling the models to account for cross-market dynamics and interdependencies, a factor that significantly boosted forecasting potential and addressed a gap in existing literature. The meticulous feature selection process, guided by Random Forest feature importance, also played a crucial role, particularly by highlighting the predictive power of lagged Bitcoin prices and technical indicators like EMAs, and justifiably leading to the exclusion of features with minimal impact on predictive performance, such as the sentiment scores in the final models.

8.2 Model Performance Overview

The comparative analysis of the implemented models on the unseen test set provided clear evidence of their respective strengths and weaknesses in the context of Bitcoin price forecasting. The LSTM model distinguished itself as the top performer, achieving a remarkable R^2 of **0.89** and a low Mean Absolute Error (MAE) of **4284.69**. Its ability to explain a significant portion of the variance in the test data highlights its robustness and suitability for predicting highly volatile assets. The Gated Recurrent Unit (GRU) model also exhibited strong performance, securing an R^2 of **0.79** and an MAE of **6515.65**. While slightly less accurate than the LSTM, the GRU model offers a compelling balance of predictive power and potentially lower computational complexity. In contrast, the ensemble methods, Random Forest and XGBoost, despite demonstrating strong performance on validation data (R^2 of **0.95** and **0.93** respectively), suffered significant drops in performance on the test set (R^2 of **-0.51** and **-0.55**), indicating issues with generalization and overfitting. Traditional time-series models such as ARIMA (R^2 **-0.95**, MAE **18662.27**) and Prophet (R^2 **0.13**, MAE **13061.58**) proved ill-suited for the non-linear and unpredictable nature of Bitcoin prices, yielding poor performance metrics on the test set.

8.3 Implications for Real-World Application

The superior performance of the LSTM model, characterized by its high accuracy and relatively low error rates (MAPE of **4.95%** and Directional Accuracy of **53.31%**), supports its validity for potential real-world applications. For traders and financial institutions, models with such predictive capabilities can inform decision-making processes, potentially enabling the development of more sophisticated trading strategies, improved risk management, and the ability to anticipate market movements with greater confidence. The insights gained from the inclusion of multi-cryptocurrency data further enhance the strategic value of the framework, allowing stakeholders to consider the broader market ecosystem and its influence on Bitcoin's price dynamics.

8.4 Limitations and Future Directions

Despite the project's successes, several limitations were identified that present avenues for future research. The exclusion of sentiment data, while justified by feature importance in this study, warrants further investigation with potentially more granular or alternative sentiment metrics to fully assess its predictive value. The models, including the top-performing LSTM, consistently struggled with accurately predicting extreme price fluctuations due to the inherent rarity of these events in historical data; exploring techniques for handling outliers or incorporating synthetic data could address this. Expanding the scope beyond Bitcoin to include a wider range of macroeconomic indicators, on-chain analytics, and alternative data sources could provide a more holistic view of market drivers and potentially improve model robustness. Furthermore, computational constraints limited the scale and complexity of the models and datasets used; leveraging distributed computing resources could enable training on larger datasets and exploring deeper architectures. Future work could also focus on developing real-time prediction systems, exploring more advanced models like Transformers, or investigating reinforcement learning for optimizing trading strategies based on model outputs.

8.5 Final Thoughts

In summary, this project successfully demonstrated the efficacy of using advanced machine learning techniques, particularly LSTM and GRU, coupled with multi-cryptocurrency data and a data-driven feature selection process, for forecasting Bitcoin prices. While challenges remain, particularly in predicting extreme volatility, the framework developed provides a robust foundation for future research and holds significant potential for practical application in navigating the complexities of the cryptocurrency market. By addressing the identified limitations and continuing to explore innovative modeling approaches, the accuracy and reliability of Bitcoin price predictions can be further enhanced, contributing valuable tools for stakeholders in this evolving financial landscape.

References

Aidoo, D. and Ababio, K.A. (2023) *Modeling Bitcoin prices and returns using ARIMA model*. International Journal of Innovation and Development, Special Edition (December), pp.62–64. Available at: https://ijid.kstu.edu.gh/index.php/ijid/article/download/22/18/69 (Accessed: 7 May 2025).

Bakar, N. and Rosbi, S. (2017) *Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate*. International Journal of Advanced Engineering Research and Science, 4(5), pp.123–129.

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014) *Learning phrase representations using RNN encoder-decoder for statistical machine translation*. arXiv preprint arXiv:1406.1078.

CoinMarketCap (2025) *Cryptocurrency market data*. Available at: https://coinmarketcap.com (Accessed: 7 May 2025).

Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2018) *BERT: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.

Economic Times (2025) *Understanding the role of Bitcoin in modern investment portfolios*. Available at: https://economictimes.com/markets/cryptocurrency/understanding-the-role-of-bitcoin-in-modern-investment-portfolios/articleshow/116939447.cms (Accessed: 7 May 2025).

ElKulako, A. (2022) *CryptoBERT: A pre-trained model for cryptocurrency sentiment analysis*. Journal of Financial NLP, 12(3), pp.45–52.

Fama, E.F. (1970) *Efficient capital markets: A review of theory and empirical work*. The Journal of Finance, 25(2), pp.383–417.

Hochreiter, S. and Schmidhuber, J. (1997) *Long short-term memory*. Neural Computation, 9(8), pp.1735–1780.

Kraaijeveld, O. and De Smedt, J. (2020) *The predictive power of public Twitter sentiment for forecasting cryptocurrency prices*. Journal of International Financial Markets, Institutions and Money, 65, 101112.

Kuizinienė, D., Dagienė, E. and Kašauskas, M. (2019) *Machine learning techniques for cryptocurrency price prediction: A comparative study*. Applied Economics Letters, 26(15), pp.1234–1239.

Kulakowski, M. and Frasincar, F. (2023) *Sentiment classification of cryptocurrency-related social media posts*. IEEE Intelligent Systems, 38(4), pp.5–9. https://doi.org/10.1109/MIS.2023.3283170.

Li, Y., Zhang, Y., Wang, J. and Liu, Y. (2021) *Integrating Twitter sentiment with LSTM for Bitcoin price prediction*. IEEE Transactions on Computational Social Systems, 8(4), pp.921–930.

Modi, P.D., Arshi, K., Kunz, P.J. and Zoubir, A.M. (2023) *A data-driven deep learning approach for Bitcoin price forecasting*. arXiv preprint. Available at: https://arxiv.org/pdf/2311.06280.pdf (Accessed: 7 May 2025).

Mudrex (2025) *Politics and Bitcoin: How politics influences Bitcoin*. Available at: https://mudrex.com/learn/politics-and-bitcoin/ (Accessed: 7 May 2025).

Nakamoto, S. (2008) *Bitcoin: A peer-to-peer electronic cash system*. Available at: https://bitcoin.org/bitcoin.pdf (Accessed: 7 May 2025).

OSL (2025) *Why Bitcoin is the future of financial independence*. Available at: https://osl.com/academy/article/why-bitcoin-is-the-future-of-financial-independence (Accessed: 7 May 2025).

OSL (2025) *What makes Bitcoin's price go up?* Available at: https://osl.com/academy/article/what-makes-bitcoins-price-go-up (Accessed: 7 May 2025).

Pepperstone (2024) *What is Bitcoin trading and how do you trade it?* Available at: https://pepperstone.com/en-eu/learn-to-trade/trading-guides/how-to-trade-bitcoin/ (Accessed: 7 May 2025).

Shiller, R.J. (2003) *From efficient markets theory to behavioral finance*. Journal of Economic Perspectives, 17(1), pp.83–104.

Siami-Namini, S., Tavakoli, N. and Siami Namin, A. (2019) *A comparison of LSTM for time series forecasting*. Procedia Computer Science, 159, pp.1213–1222.

Souza, S.S. and Silva, J.E. (2024) *Application of LSTM recurrent neural networks for Bitcoin price prediction*. Delos: Desenvolvimento Local Sustentável, 15(45). Available at: https://ojs.revistadelos.com/ojs/index.php/delos/article/view/2602 (Accessed: 7 May 2025).

Tripathi, R. and Sharma, S. (2023) *Sentiment analysis for cryptocurrency price prediction using BERT and influencer weighting*. Computational Finance Journal, 19(2), pp.67–75.

Appendix – A

Github Repository - https://github.com/Fizryfu/Capstone/tree/main/Capstone%20Project