BYTeminds' CryptoStrategist: Enhancing Bitcoin Profits through Smart Predictive Models

Sathark Sharma, Manan Chugh , Piyush Sinha , Mishita Jain :-Netaji subhas University of Technology , New Delhi ,India
Email:{sarthak062004@gmail.com, mananchugh2004@gmail.com, piyushsinha19807@gmail.com, jainmishita1977@gmail.com}

Abstract—

The paced changes happening in the cryptocurrency market, with Bitcoin offer great opportunities and challenges for investors. In this project led by the BYTeminds team we use machine learning techniques to maximize profits from Bitcoin by making predictions. Our research focuses on analyzing data using algorithms and creating a predictive model that can optimize trading strategies.

We start by discussing the difficulties faced in cryptocurrency trading and emphasize the importance of having tools for making predictions. The project then dives into an exploration of machine learning algorithms that are well suited for the dynamic nature of the crypto market. Our model is trained using a dataset and historical trends.

The results demonstrate how our approach effectively predicts changes in Bitcoin prices and helps optimize trading decisions to maximize profits. We also discuss insights gained from our models' performance. Identify areas where improvements can be made. Additionally, our report discusses the implications of our findings for investors and the cryptocurrency ecosystem.

Our project contributes to the growing knowledge base at the intersection of machine learning and cryptocurrency trading providing insights and strategies, for individuals and institutions navigating through this crypto market.

Introduction--

Analyzing market trends through historical data poses a formidable challenge, especially when dealing with sizable datasets. Conventional statistical methods often falter in the face of such massive information reservoirs. In this area predicting next market trend is becoming increasingly important. In this report, we detail our approach to transcend traditional statistical limitations and we used Deep Learning to unlock insights, forecast trends, and optimize trading strategies in the ever-evolving cryptocurrency market.

Dataset:

1.Training, Validation, and Test Data (Jan 1, 2018, to Jan 31,2022)

This period is used for training, validation, and testing our models.

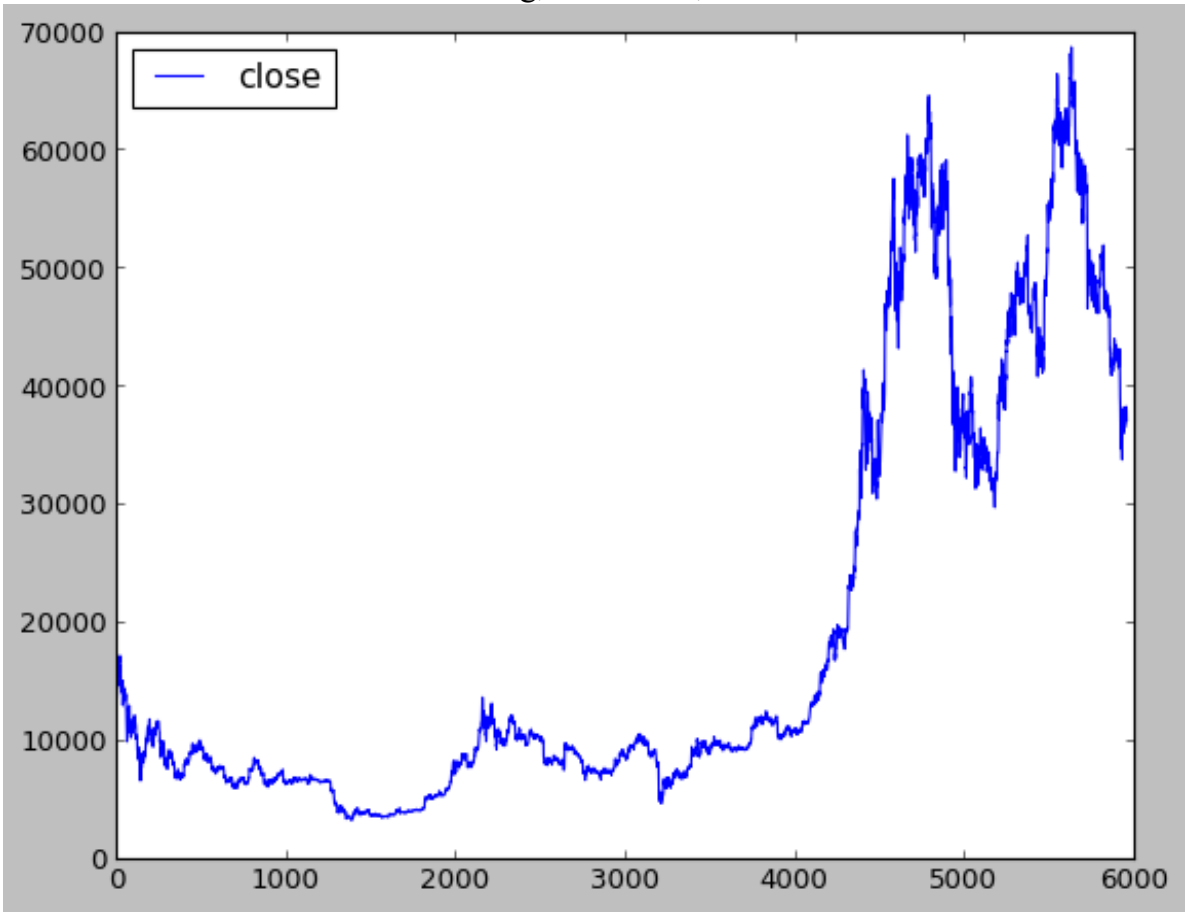
2. Out-of-Sample 1 (Feb 1, 2022, to Dec 31, 2022)

- A separate dataset is used to assess our model's performance in an out-of-sample scenario.

3. Out-of-Sample 2 (Jan 1, 2023, to Dec 31, 2023)

- Another dataset is used to evaluate our model's generalization capabilities further.

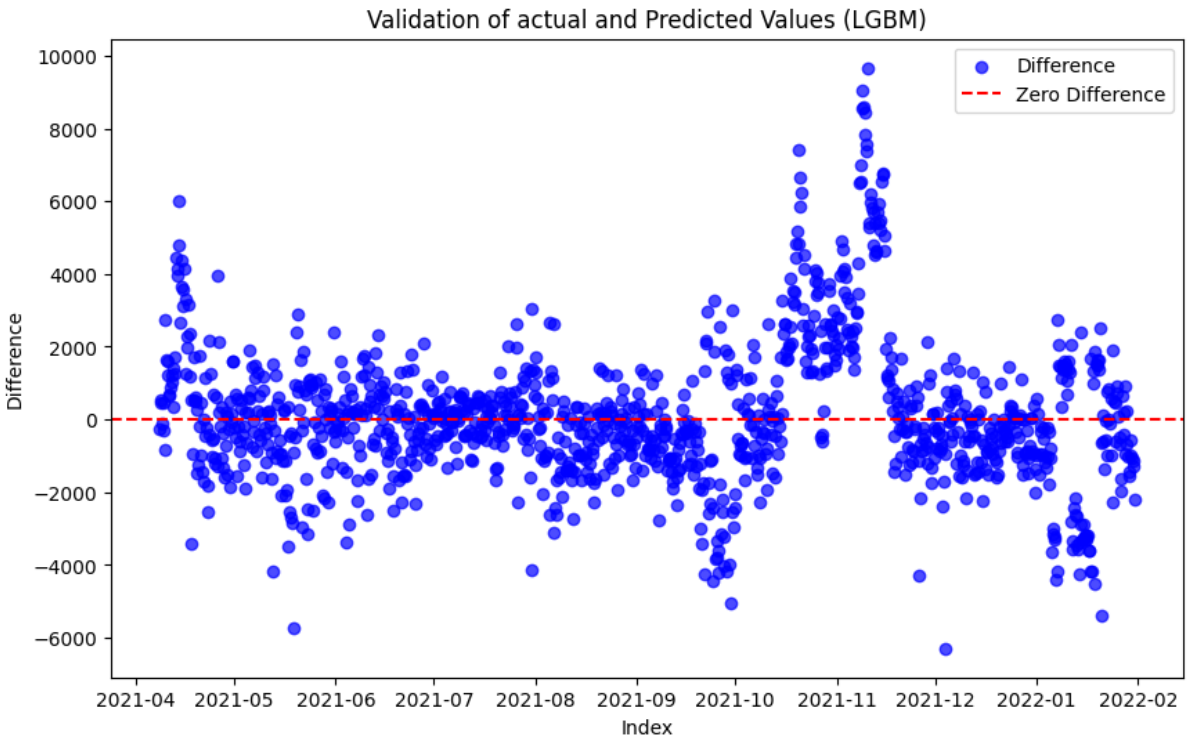
Plot of the Training, Validation, and Test Data



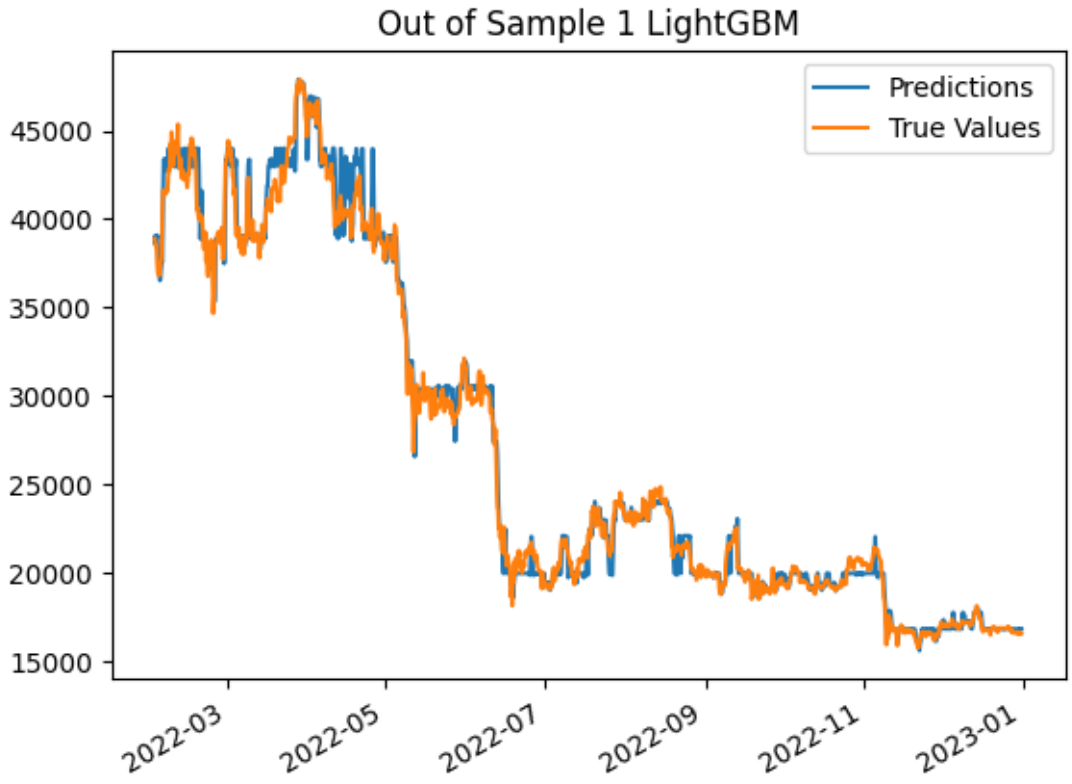
Models:

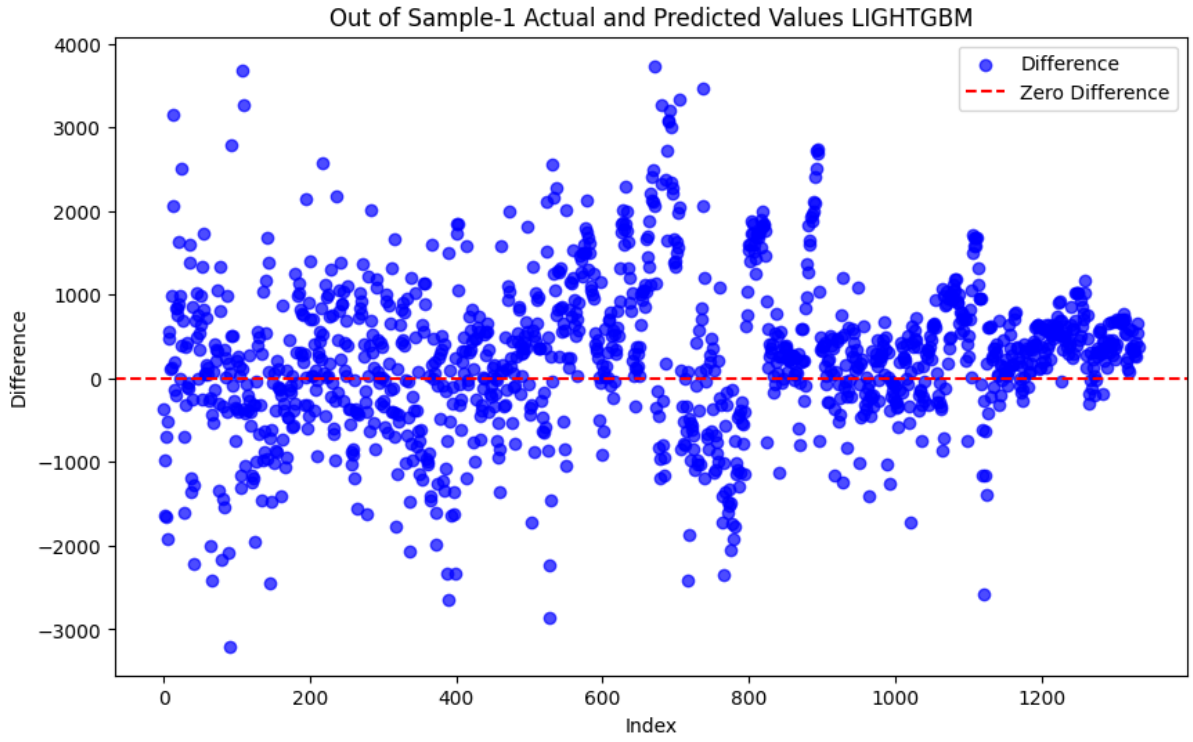
1: LGBMRegressor
2: XGBRegressor
3: CatBoostRegressor
4.CNN

LGBM Regressor:

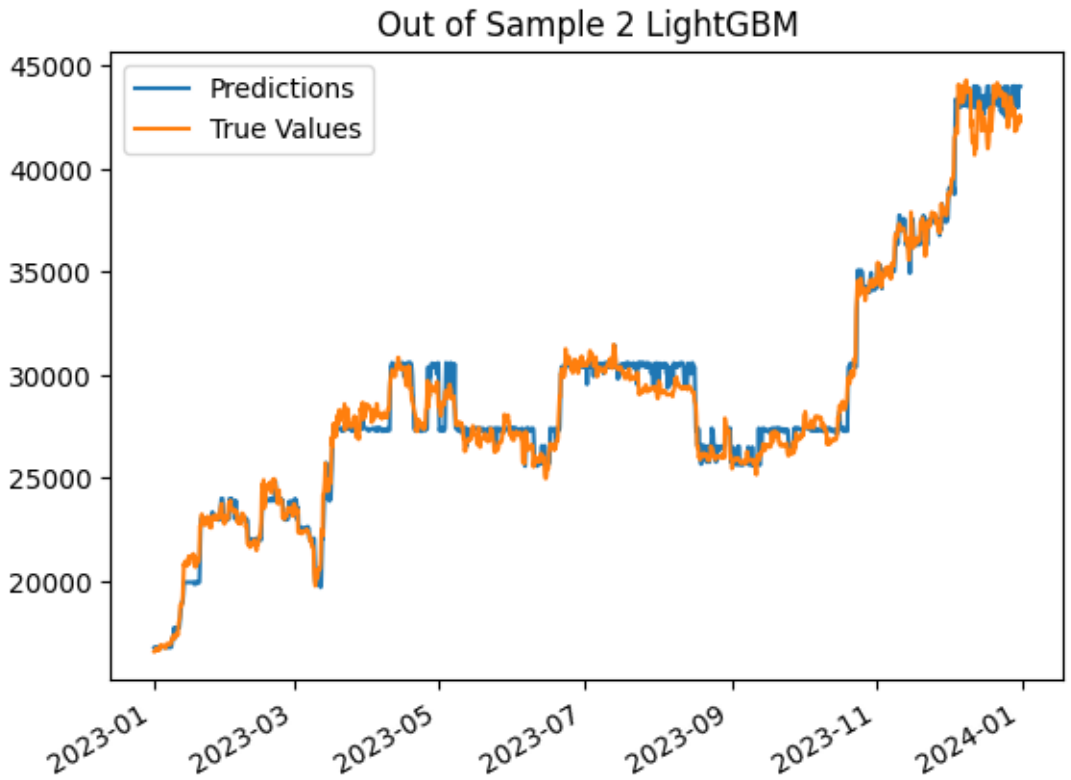


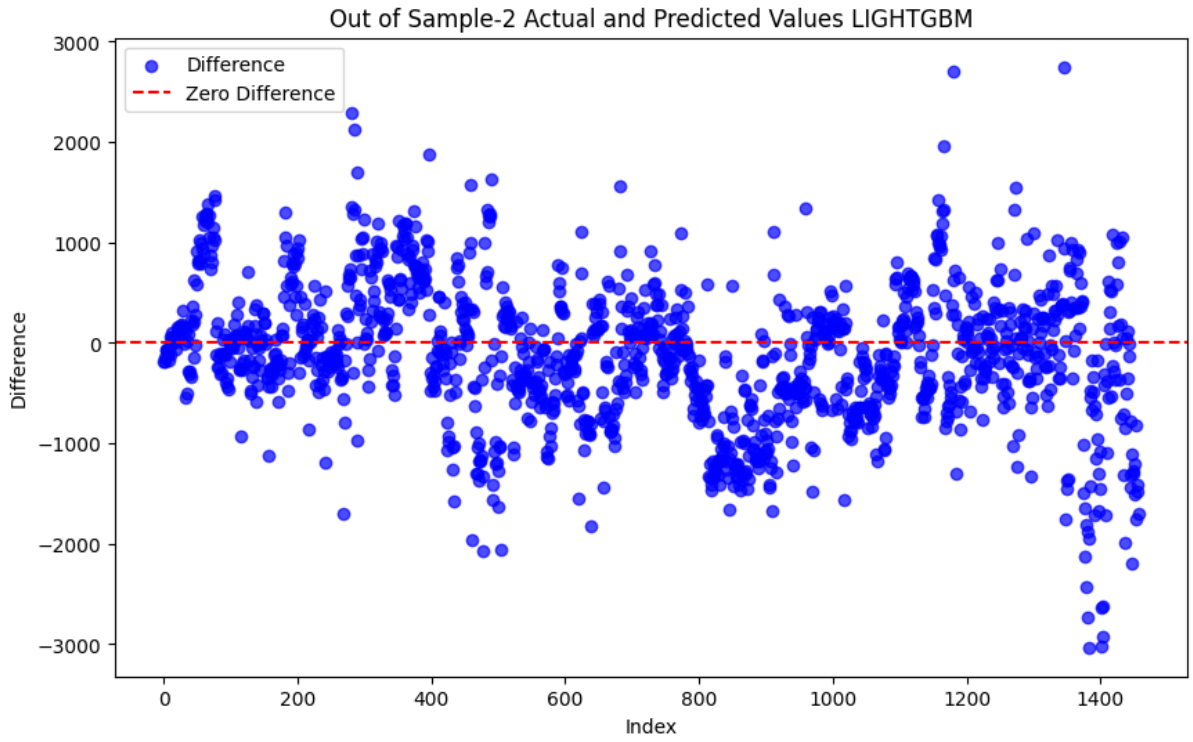
Testing LightGBM on Out-of-Sample 1 (Feb 1, 2022, to Dec 31, 2022)





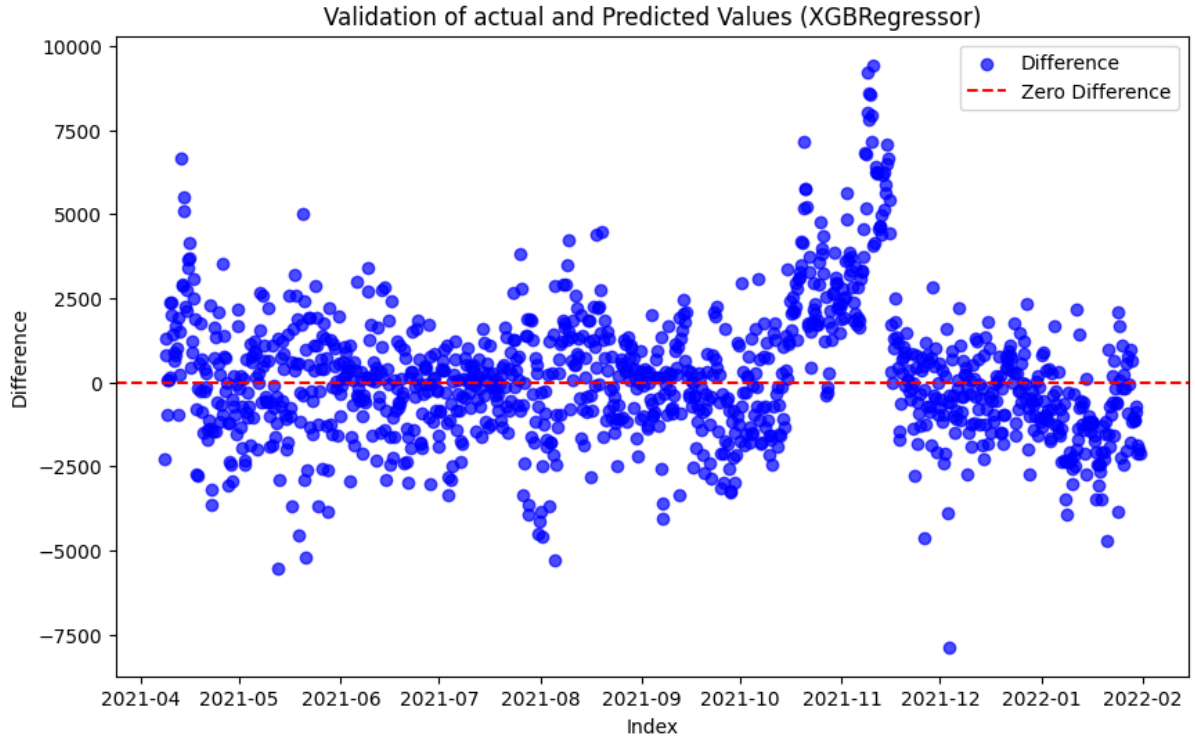
Testing LightGBM on Out-of-Sample 2 (Jan 1, 2023, to Dec 31, 2023)



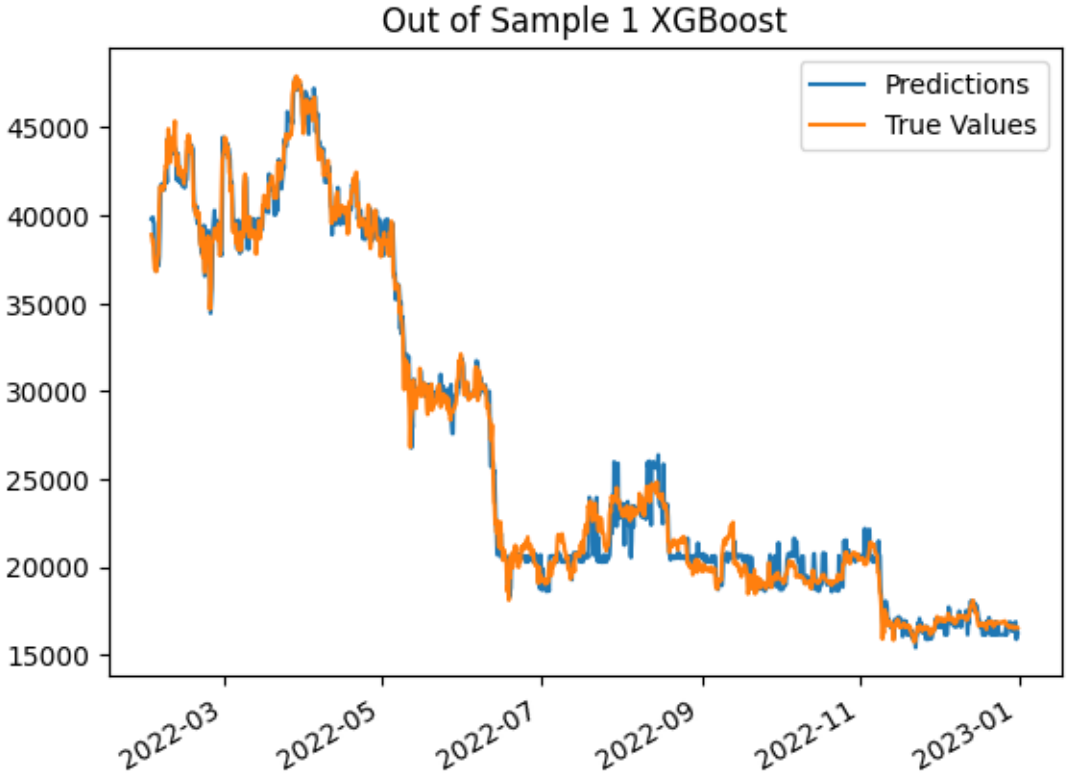

Out of Sample-2 Actual and Predicted Values LIGHTGBM
Difference
Zero Difference
3000
2000
1000
0
−1000
−2000
−3000
Difference
0
200
400
600
800
1000
1200
1400
Index

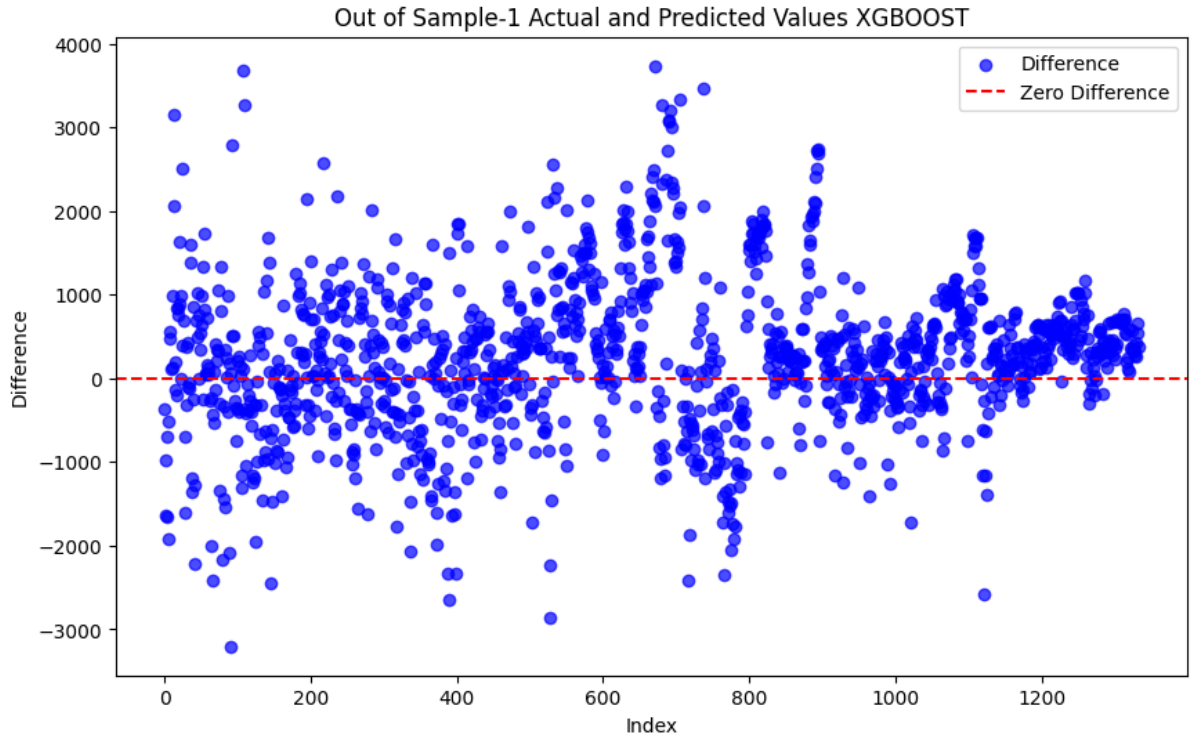
2. XGBRegressor

This graph shows the difference in the value of the testing and training data

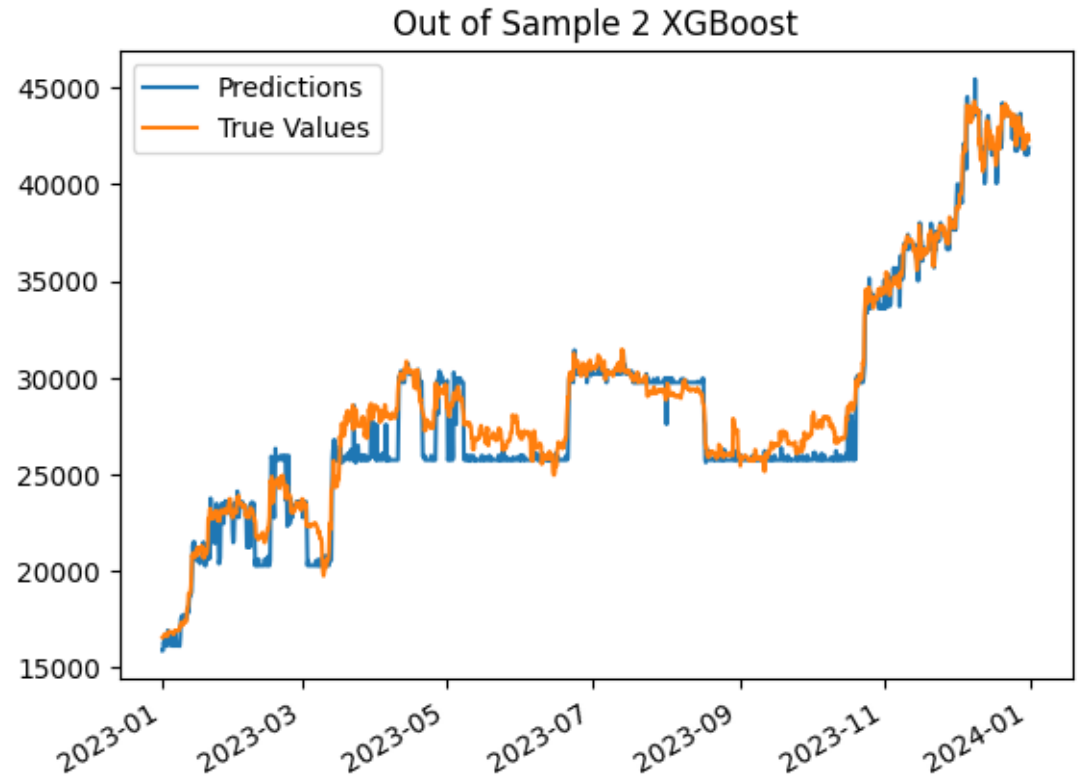


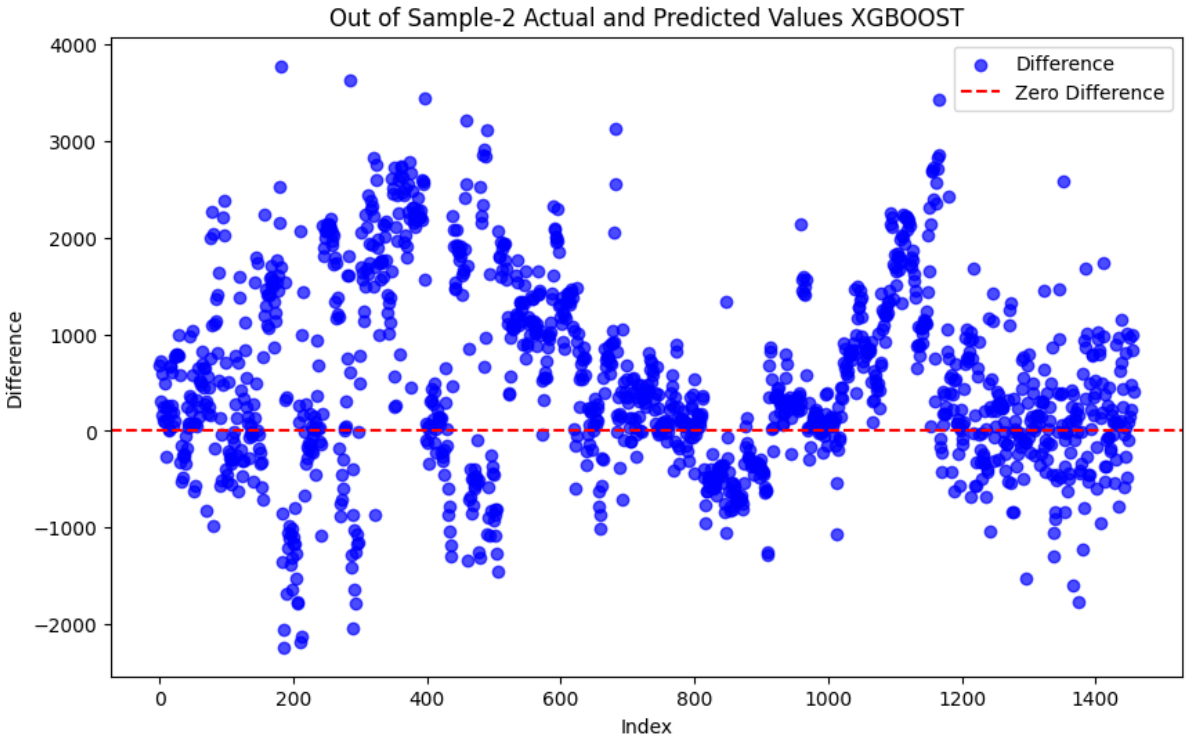
Testing XGBRegressor on Out-of-Sample 1 (Feb 1, 2022, to Dec 31, 2022)





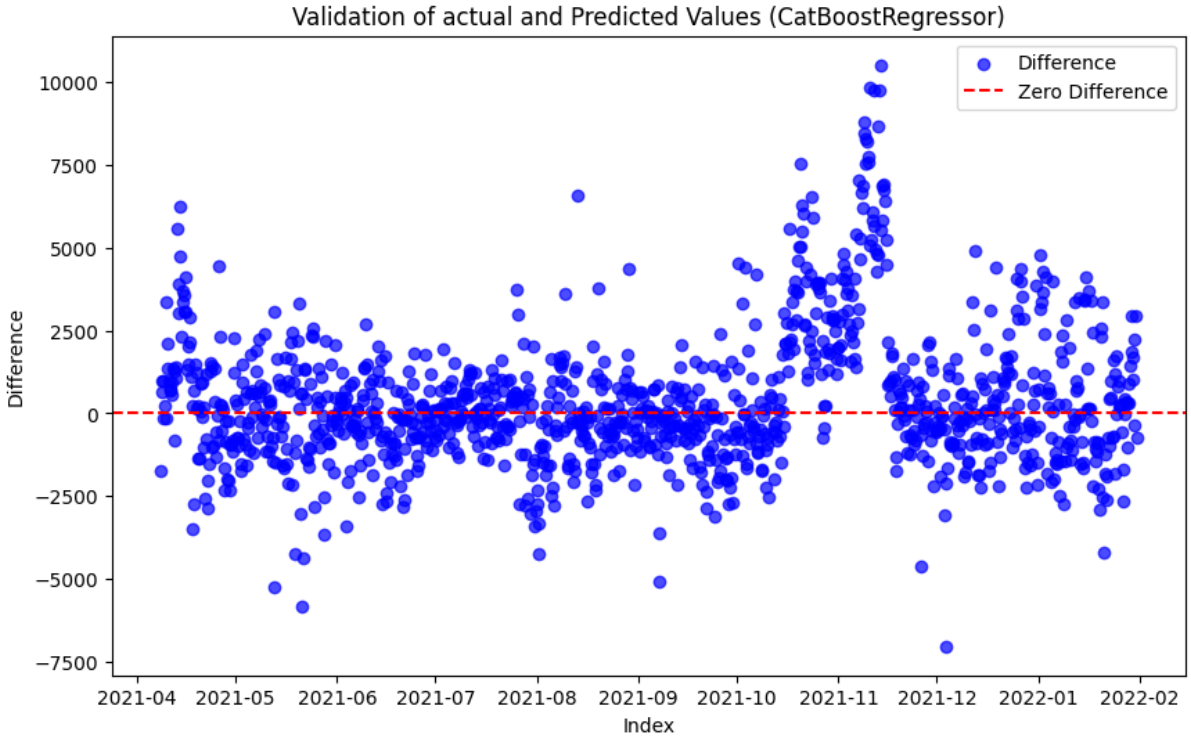
Testing XGBRegressor on Out-of-Sample 2 (Jan 1, 2023, to Dec 31, 2023)



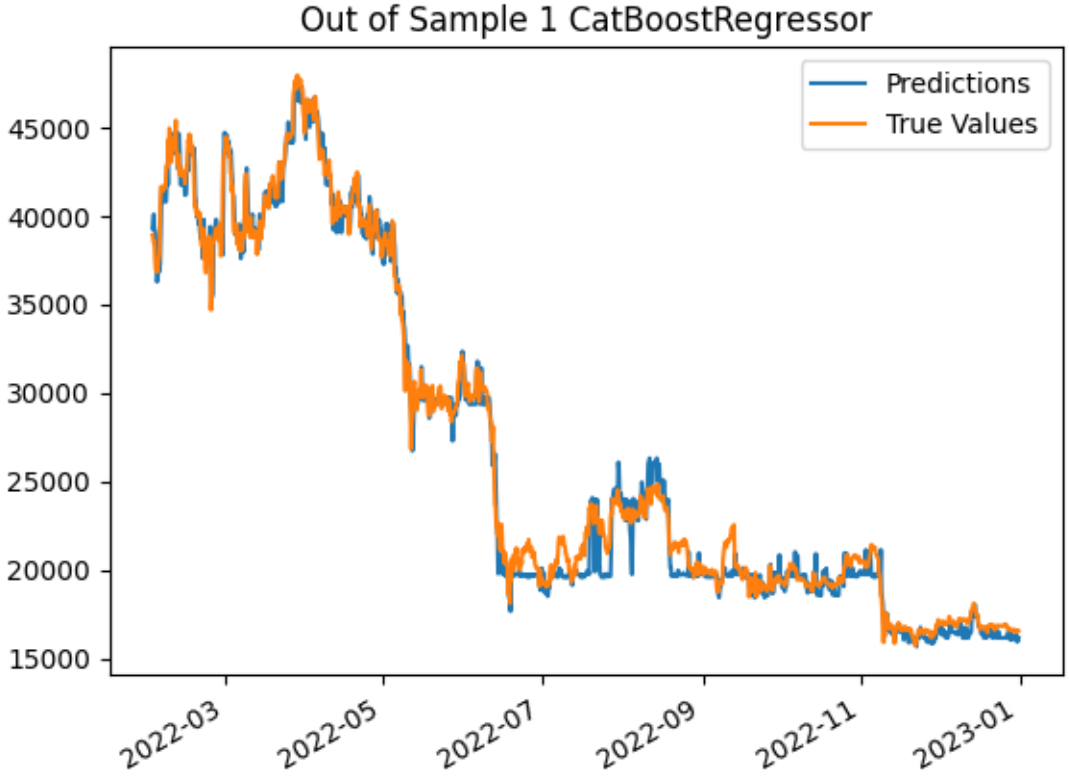

Out of Sample-2 Actual and Predicted Values XGBOOST
Difference
Zero Difference
Difference
Index
4000
3000
2000
1000
0
−1000
−2000
0
200
400
600
800
1000
1200
1400

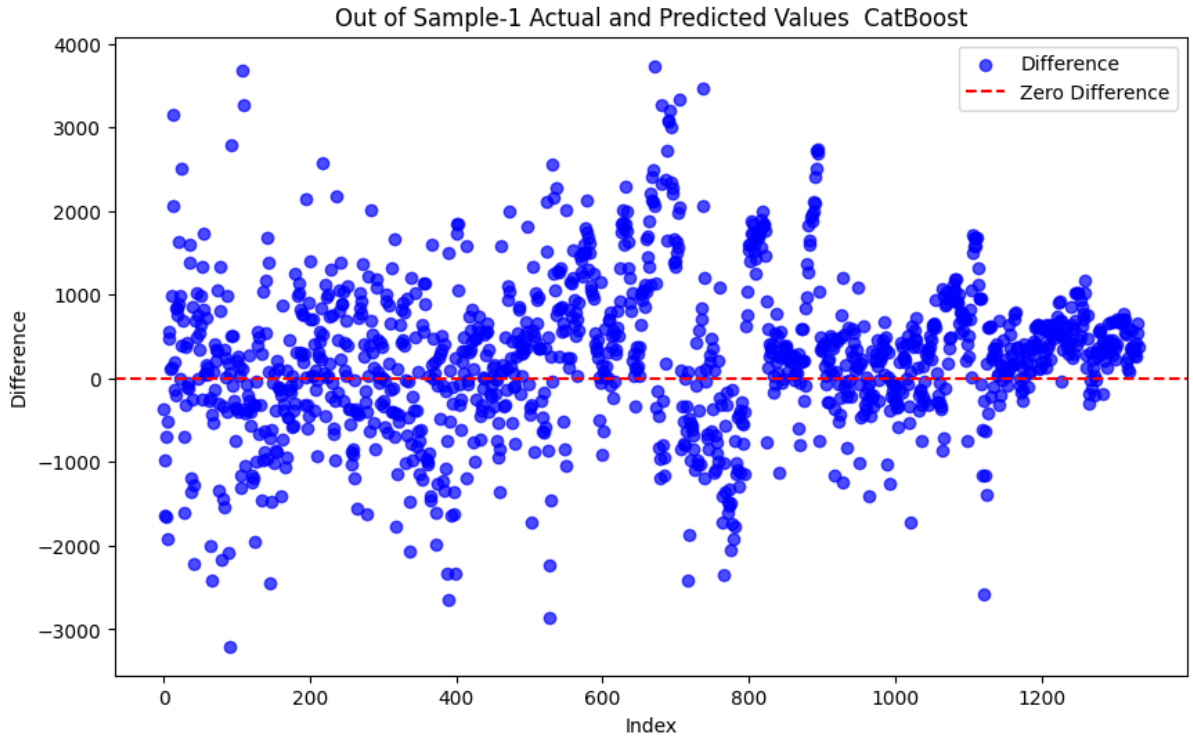
3: CatBoostRegressor

This graph shows the difference in the value of the testing and training data

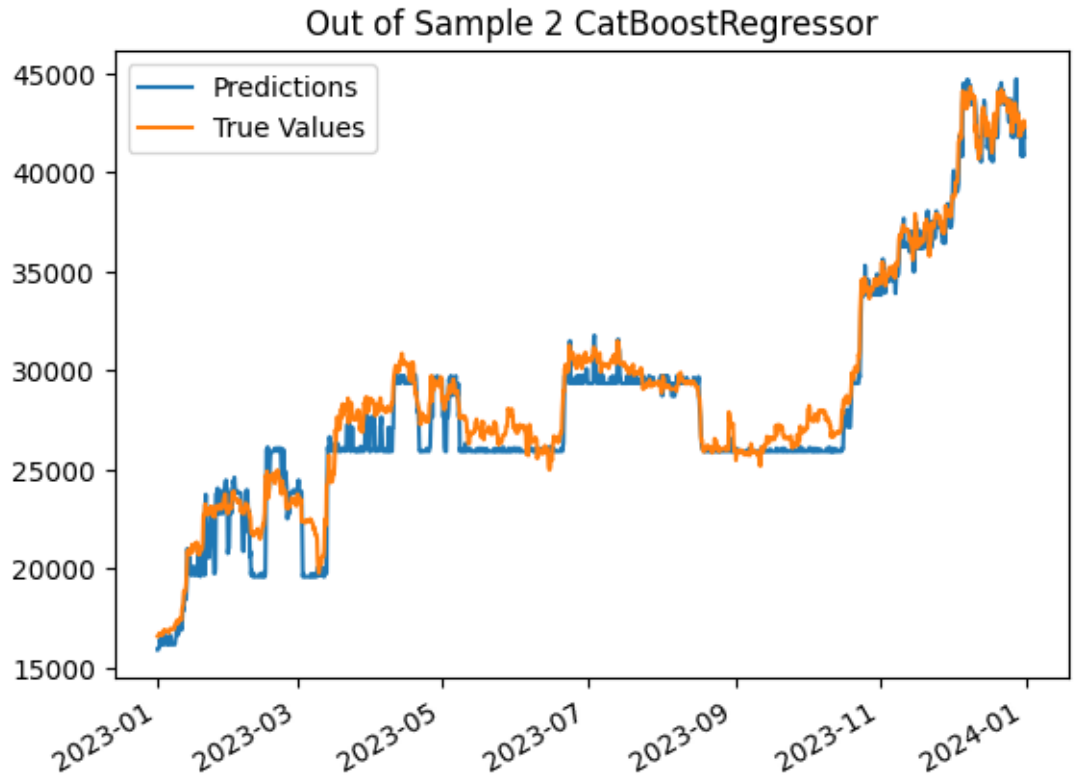


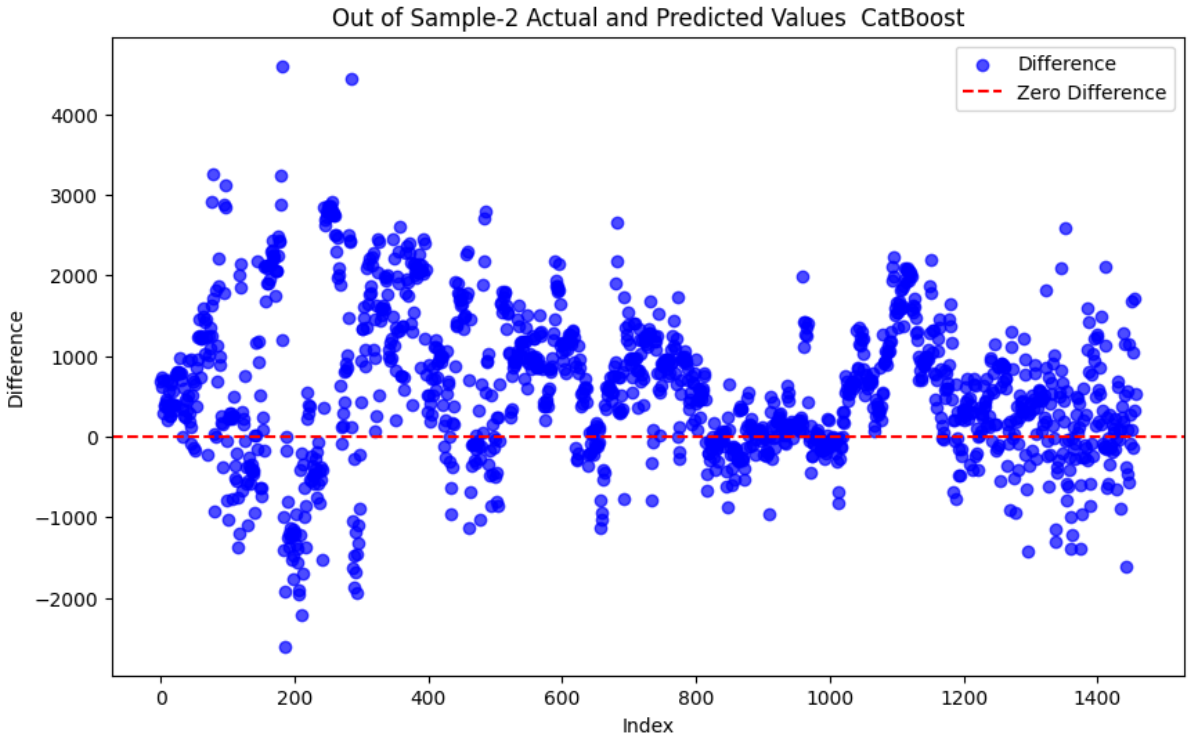
Testing CatBoostRegressor on Out-of-Sample 1 (Feb 1, 2022, to Dec 31, 2022)





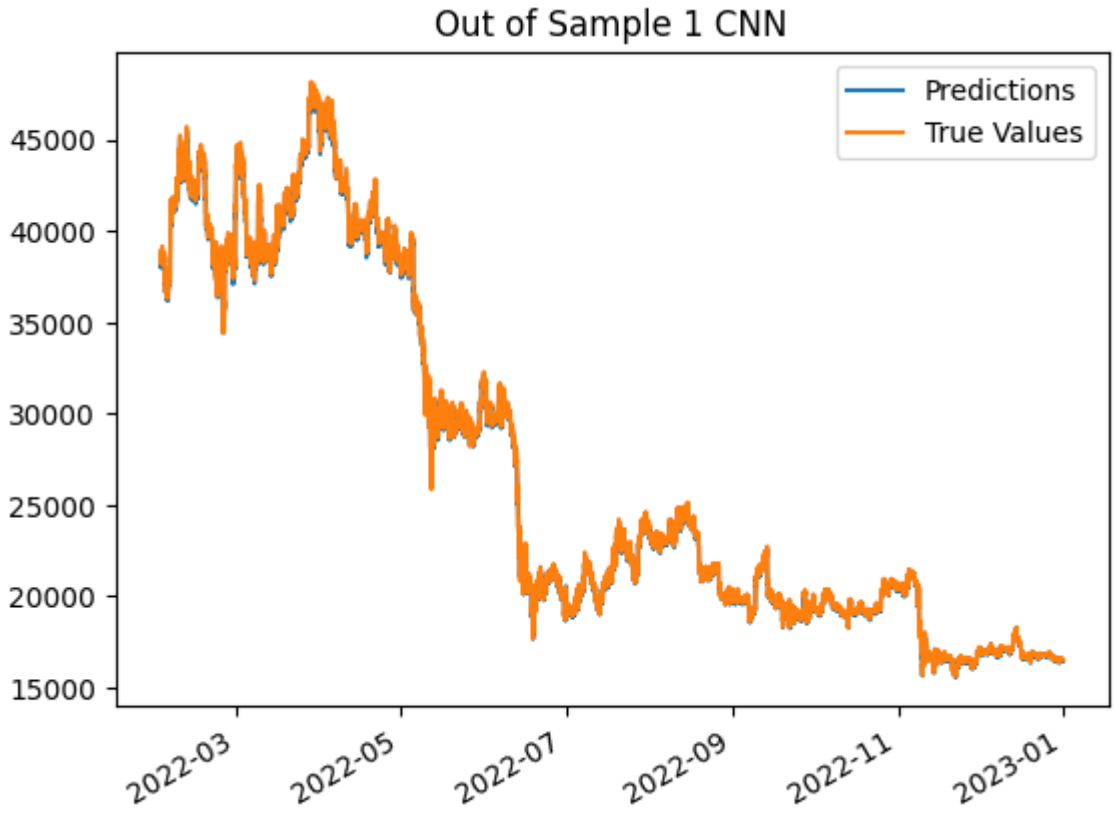
Testing CatBoostRegressor on Out-of-Sample 2 (Jan 1, 2023, to Dec 31, 2023)

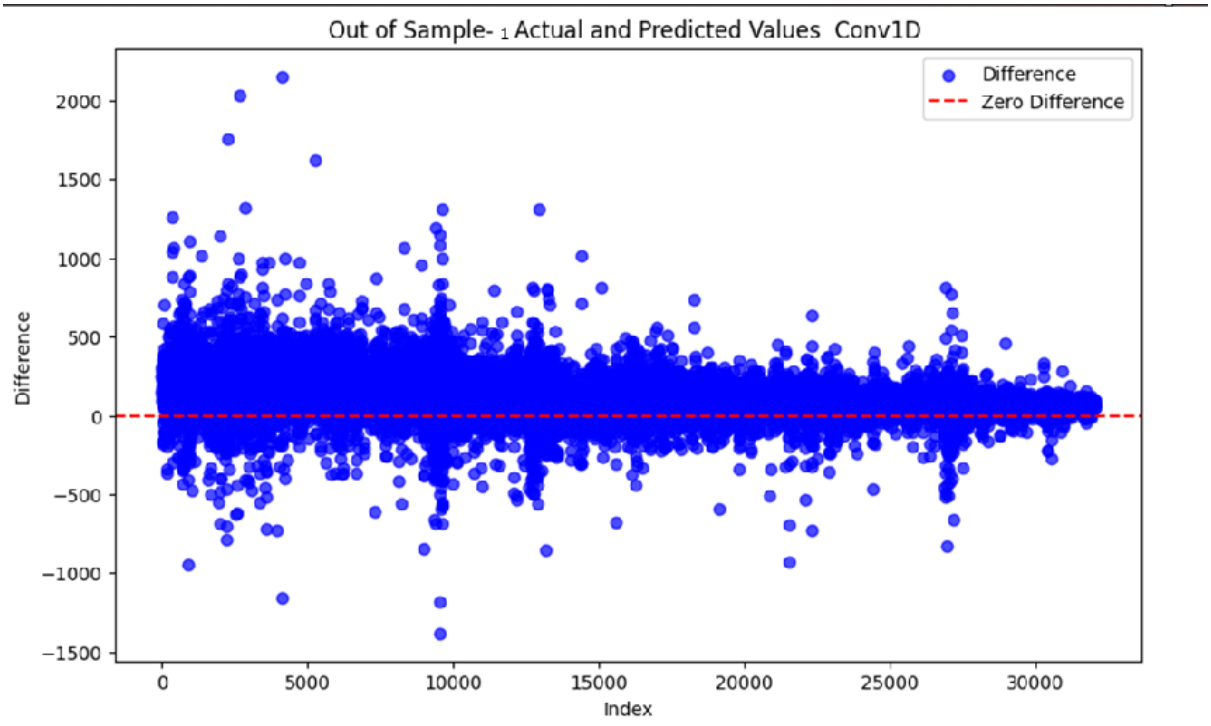



Out of Sample-2 Actual and Predicted Values CatBoost
Difference
Zero Difference
4000
3000
2000
1000
0
−1000
−2000
Difference
0
200
400
600
800
1000
1200
1400
Index

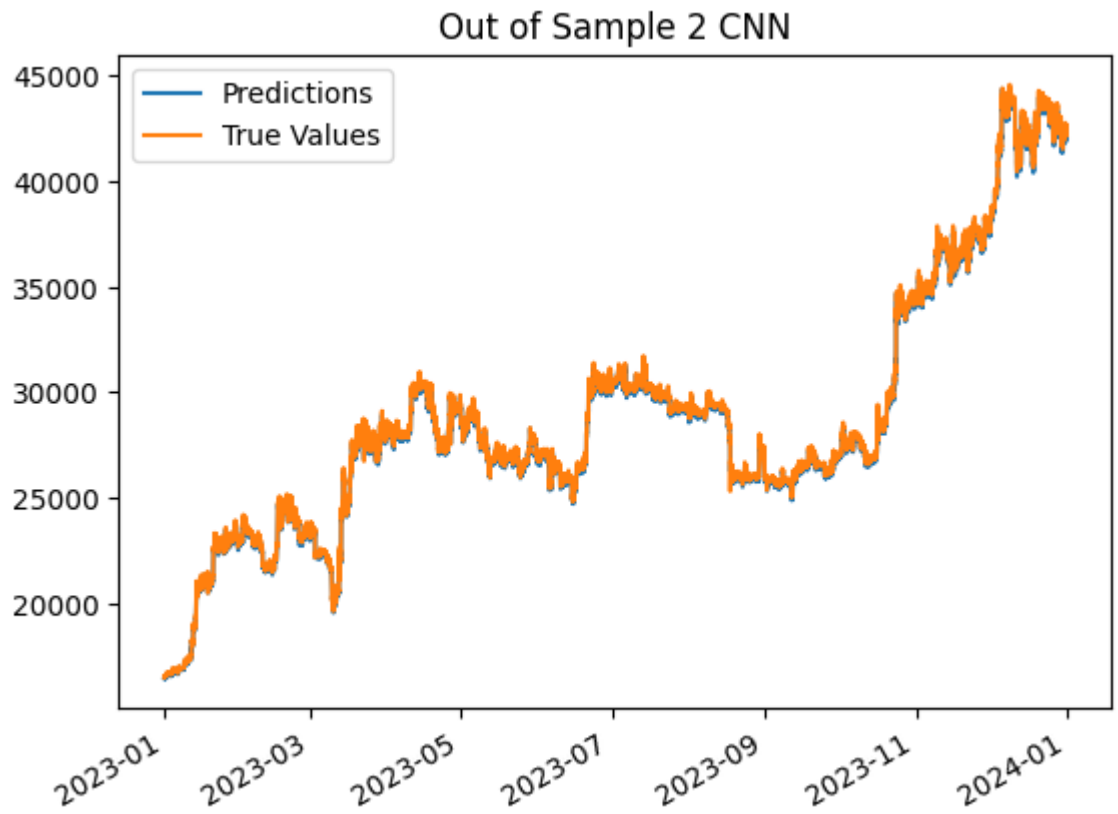
CNN

Testing CNN on Out-of-Sample 1 (Feb 1, 2022, to Dec 31, 2022)





Testing CNN on Out-of-Sample 2 (Jan 1, 2023, to Dec 31, 2023)



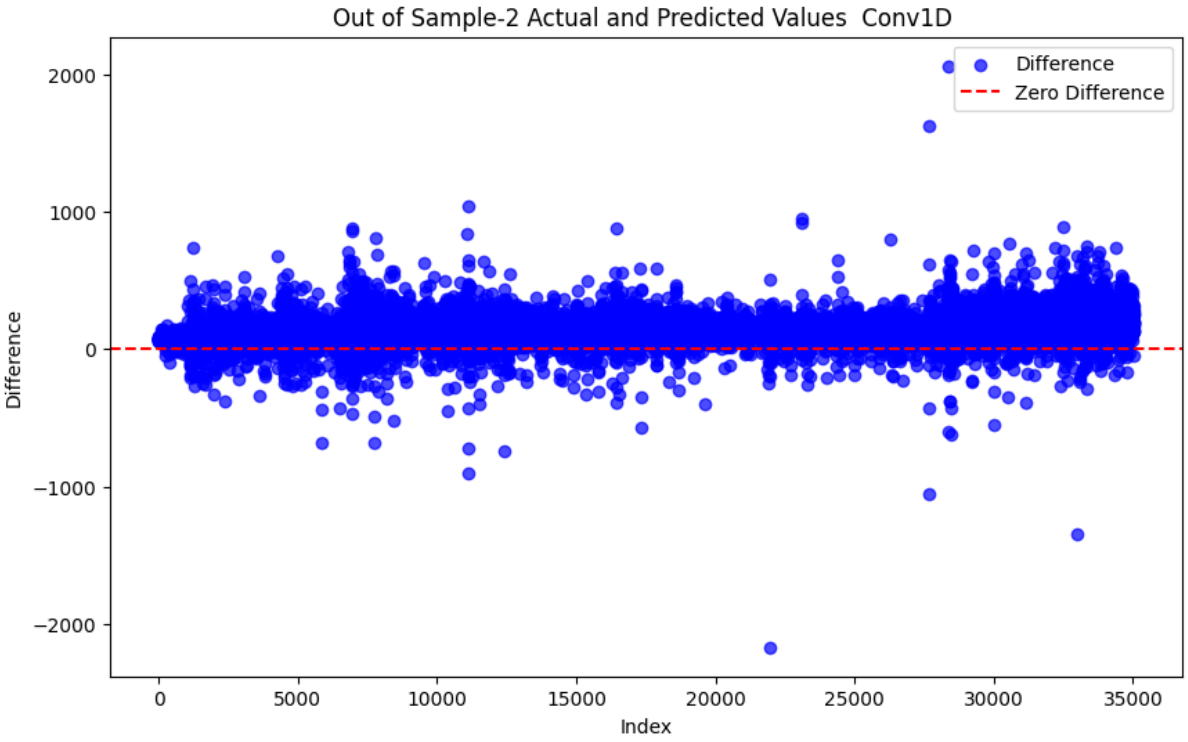


Table : Mean Absolute Error

MODEL	Mean Absolute Error {mae}
i. LGBMRegressor	a. 1342.5852114855377
ii. XGBRegressor	b. 1454.4816446885488
iii. CatBoostRegressor	c. 1454.4816446885488
iv. CNN	d. 688.2727356858222

Model Selection and Justification: Convolutional Neural Network (CNN)

After considering and evaluating various models for their suitability in our regression task, the Convolutional Neural Network (CNN) was chosen for its unique architecture and specific advantages that align with the characteristics of our data.

1. **Temporal Pattern Recognition:** While traditionally associated with image and grid-based data, CNNs can also be effective in capturing temporal patterns within sequential data. In the context of Bitcoin price prediction, the historical price data can be treated as a sequence, and the CNN's convolutional layers can be adapted to recognize meaningful temporal patterns and relationships.

2. **Non-linear Feature Extraction:** CNNs are known for their ability to automatically learn hierarchical representations of features, including non-linear relationships. In cryptocurrency markets, where price movements are influenced by a myriad of factors with complex interactions, the CNN's capacity to extract intricate and non-linear features can be advantageous for accurate prediction.

3. **Localized Dependencies:** The CNN's convolutional layers are designed to exploit local patterns and dependencies within the data. In the case of Bitcoin price prediction, where short-term fluctuations and localized trends can significantly impact the overall trajectory, the CNN's focus on capturing such local dependencies can contribute to improved predictive accuracy.

4. **Performance Metrics:** As evidenced by the achieved Mean Absolute Error (MAE) of 688.2727356858222, the CNN model demonstrated superior predictive accuracy compared to alternative models such as LGBMRegressor, XGBRegressor, and CatBoostRegressor. This underscores the adaptability of CNNs in capturing temporal patterns and non-linear dependencies present in our Bitcoin price dataset.

In conclusion, the inclusion of the CNN model in our regression analysis for Bitcoin price prediction was motivated by its capacity to recognize temporal patterns, extract non-linear features, and capture localized dependencies within the historical price data. The lower Mean Absolute Error achieved further supports the effectiveness of the CNN in our specific predictive task.

Strategy Overview:

1. **Trading Decision:**

   The trading decision in this strategy is based on a **mean-reversion approach**, aiming for profit from short-term price deviations. Here's a brief overview:

   a. **Position Management:**
      i. The strategy starts by managing the current position, checking if the market conditions are favourable for selling. If so, it assesses conditions for selling, including stop-loss, bullish signals, and drawdown management.
   b. **Buying Conditions:**
      i. If not currently holding a position, the strategy looks for opportunities to buy. It considers predicted prices, recent price movements, and trend indicators to identify potential entry points.
   c. **Risk Management:**
      i. Risk is managed through the implementation of stop-loss and controlling the maximum loss.
      ii. Transaction costs are factored in for both buying and selling to account for real-world trading expenses.

   d. **Drawdown Management:**
      i. The strategy incorporates drawdown management to prevent excessive losses, ensuring that the portfolio remains within acceptable risk levels.

2. **Risk Management:**

Risk management is a crucial aspect of any trading strategy aimed at preserving capital and minimizing potential losses. In the provided trading strategy, several risk management techniques are employed:

a. Stop-Loss:

i. Stop loss orders have been set to avoid indefinite price drop cases. In case the price drops beyond certain level, the crypto currency is sold immediately

ii. If the current price falls below a certain threshold (stop-loss price), the strategy triggers a sell order to cut losses.

b. **Peak Capture Strategy:**

i. A threshold is defined to secure profits when the market reaches a peak and starts to fall off.

ii. In order to trigger the peak capture, the peak must be greater than the buying price by a certain threshold.

c. **Drawdown Management:**

i. The strategy monitors the drawdown.

ii. A condition is set to sell if the drawdown exceeds a predefined level, preventing the portfolio from experiencing excessive losses.

3. **Performance Metrics:**

Various performance metrics are calculated to evaluate the success and risk associated with the implemented mean-reversion strategy. Here's a brief explanation of each metric:

i. **Final Portfolio Value:**
The total value of the portfolio at the end of the trading period.

ii. **Profit and Loss (PnL):**
The overall financial gain or loss from the trading strategy, calculated as the difference between the final portfolio value and the initial capital.

iii. **Cumulative Returns:**
The total percentage return on the initial capital over the trading period.

iv. **Sharpe Ratio:**
A measure of the strategy's risk-adjusted performance. It indicates the excess return per unit of risk and helps assess the efficiency of the strategy.

v. **Total Closed Trades:**
The total number of trades executed throughout the trading period.

vi. **Total Won Trades:**
The number of trades that resulted in a profit
.

vii. **Winning Percentage (%):**
The percentage of total trades that were profitable, indicating the strategy's success rate.

viii. **Max Drawdown:**
The maximum percentage decline in portfolio value from a peak to a trough. It provides insights into the strategy's risk and potential losses during unfavorable market conditions.

ix. **Risk-Reward Ratio:**

a. This helps in accessing the contrast between potential profit and loss before entering in the trading process. The ratio should be favourable to the prediction such that the risk the model is taking is justified by the profit it earns
b. Assessing this ratio helps in evaluating whether the strategy is exposing itself to an acceptable level of risk relative to the potential reward.

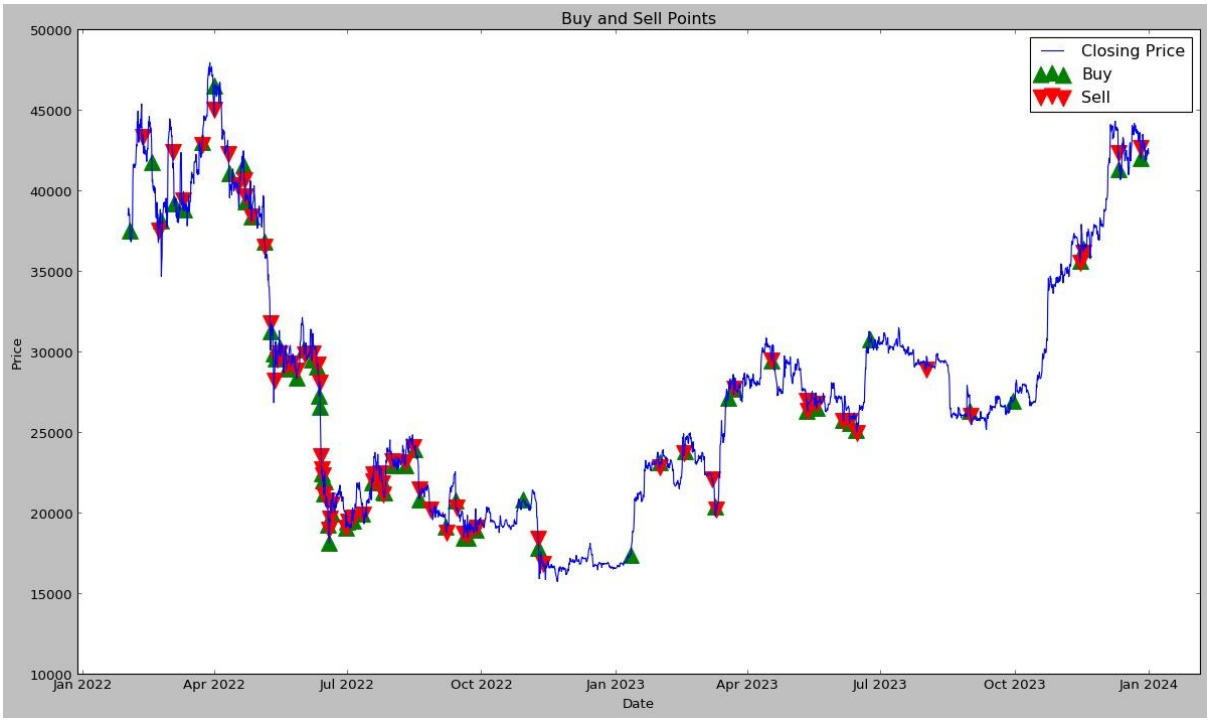
x. **Average Winning Trade:**
The average profit for each winning trade.

xi. **Average Losing Trade:**
The average loss for each losing trade.

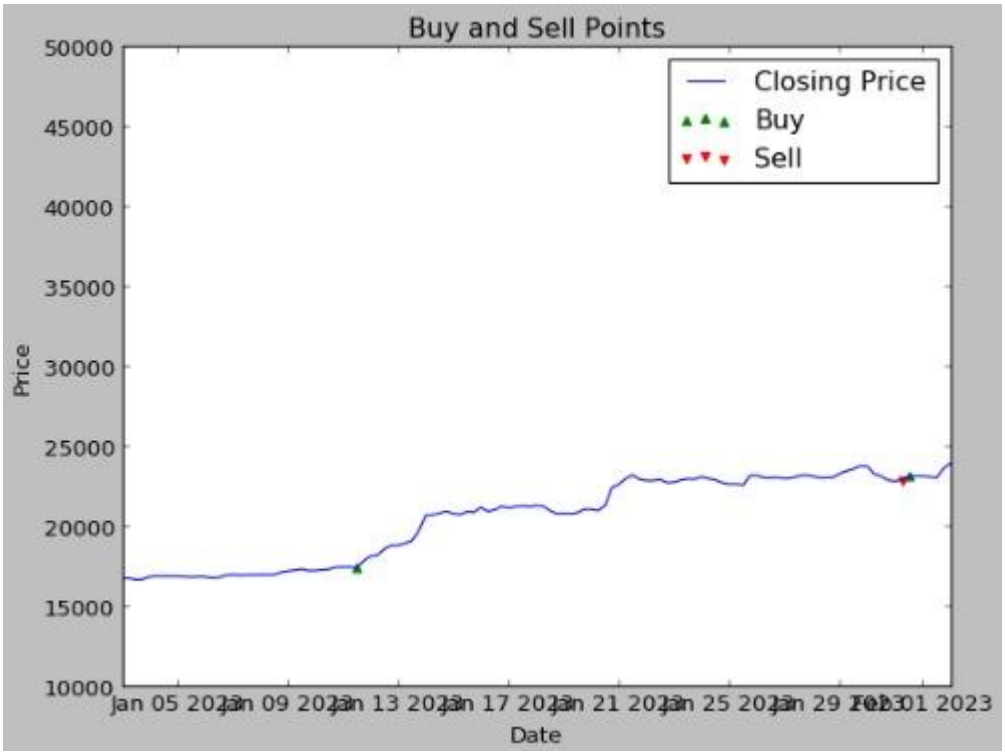
xii. **Largest Winning Trade:**
The maximum profit achieved in a single trade.

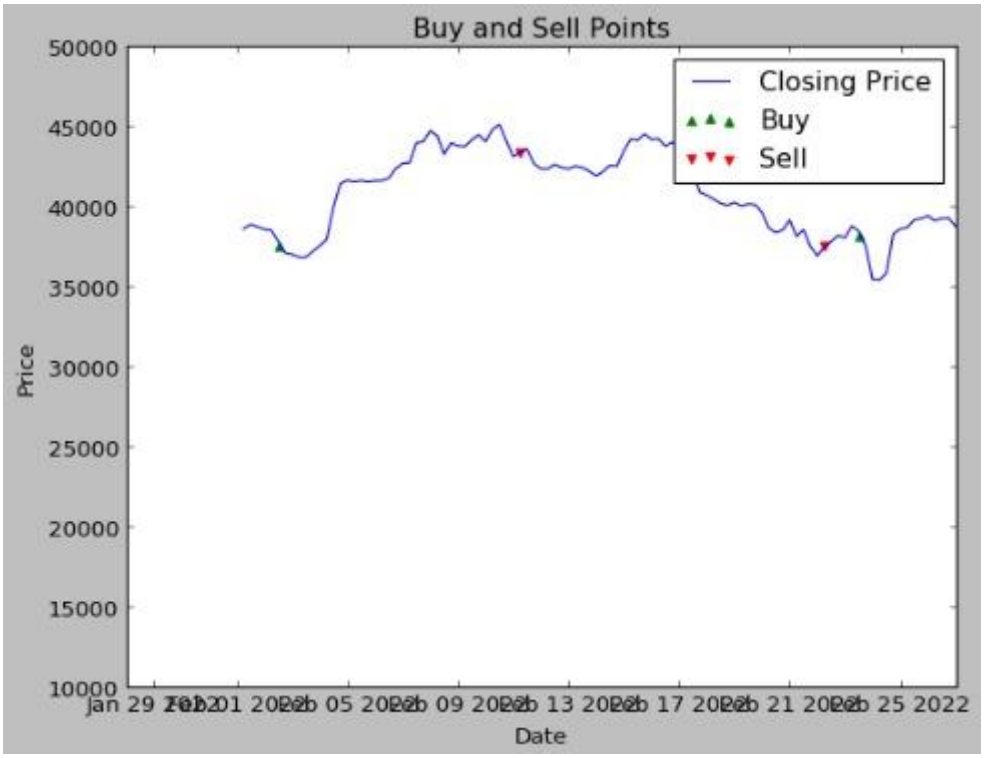
xiii. **Largest Losing Trade:**
The maximum loss incurred in a single trade.

4. **Results and Analysis:**
    - The notebook concludes with a summary of performance metrics, providing insights into the strategy's effectiveness.
    - Visualizations help in understanding the buy and sell points in the context of the historical price chart.



First capital :100000
Final Portfolio Value: 116662.72127000004
Profit and Loss: 16662.72127000004
Cumulative Returns: 0.1666272127000004
Sharpe Ratio: 1.0320492971720576
Total Closed Trades: 70
Total Won Trades: 36
Win%: 51.42857142857142
Max Drawdown: 3%
Risk Reward Ratio: 1.4042321482840108
Average Winning Trade: 1393.5213444444444
Average Losing Trade: 992.3724835294114
Largest Winning Trade: 8521.610754999994
Average Dip: 0.0348156242152614
Max Dip: 0.11952433208110744
Average Holding period: 4.757142857142857
Sortino Ratio: 14.036250482345837


Buy and Sell Points
Closing Price
Buy
Sell
Price
Date
50000
45000
40000
35000
30000
25000
20000
15000
10000
Jan 05 2023
Jan 09 2023
Jan 13 2023
Jan 17 2023
Jan 21 2023
Jan 25 2023
Jan 29 2023
Feb 01 2023


Buy and Sell Points
Closing Price
Buy
Sell
Price
Date
50000
45000
40000
35000
30000
25000
20000
15000
10000
Jan 29 2022
Feb 01 2022
Feb 05 2022
Feb 09 2022
Feb 13 2022
Feb 17 2022
Feb 21 2022
Feb 25 2022

Future research propositions

1. **Real Data Testing:** Experiment with applying machine learning to actual cryptocurrency market data.
2. **Sentiment Analysis Integration:** Explore how sentiments from social media can be used alongside machine learning for predicting market movements.
3. **Reinforcement Learning for Trading:** Check how reinforcement learning algorithms adapt trading strategies to different market conditions.
4. **Alternative Data Integration:** Look into using unconventional data sources like blockchain analytics or economic indicators to improve machine learning predictions.
5. **Handling Extreme Market Events:** Assess machine learning models' resilience during extreme market events and develop strategies for risk management.
6. **Generative Models for Data Augmentation:** Explore using generative models, like GANs, to create synthetic data and improve model training.
7. **Explainable AI for Interpretability:** Apply explainable AI techniques to make machine learning models more understandable in cryptocurrency market analysis.
8. **Scalability and Efficiency Testing:** Evaluate how well machine learning models handle larger datasets and adapt to the growing complexity of the cryptocurrency market.
9. **Feature Engineering Impact:** Study how different feature engineering techniques affect the performance of machine learning models in predicting market trends.
10. **Ensemble Learning for Accuracy:** Explore the benefits of combining multiple machine learning models through ensemble learning for improved prediction accuracy.
11. **Transferability to Different Markets:** Investigate whether machine learning models trained on one cryptocurrency market can be effective in predicting trends in other markets.