Zelta Labs, Untrade Crypto Trading Challenge

Final Report

Team 22

Inter IIT Tech Meet 12.0

1 Introduction

Our journey starts as we face the unique challenges of the BTC/USDT cryptocurrency market in the world of algorithmic trading. This field operates in a fast-paced environment where every moment can either bring opportunities or risks. As active participants, we embrace the complexity of the digital landscape and focus on creating trading techniques. These techniques aim not only to withstand market ups and downs but also to achieve returns that surpass benchmarks.

Our combined efforts result in a mix of innovation and strategic thinking. We develop and use algorithms specifically designed to navigate the intricacies of the BTC/USDT market. Beyond the potential for profits, we also prioritize the crucial aspect of managing risks, which is especially important in the unpredictable world of cryptocurrency trading.

Cryptocurrencies are always changing, requiring us to stay flexible and aware of market trends. As contributors, we present this report not just as a summary of our work but as a part of the ongoing story of innovation and expertise in BTC/USDT algorithmic trading. We hope our work sparks new ideas, encourages teamwork, and promotes methodologies where algorithms play a key role in maximizing returns and controlling risks.

2 Final Strategy

2.1 Strategy Intuition and Hypothesis

In the realm of BTC/USDT trading, where trends are prominent and market volatility is high, the problem at hand necessitates a pursuit of a risk-to-reward ratio that leans towards high rewards and low risk. The emphasis lies on identifying scenarios characterized by robust trend strength and minimal volatility. The assurance of a strong trend translates to favorable returns, while the commitment to low volatility mitigates risk. The strategy for long and short trades involves detecting trend reversals using Moving Average Convergence Divergence (MACD) and confirming trend direction with the 200-period Exponential Moving Average (EMA). To address overbought or oversold market conditions during a reversal, Relative Strength Index (RSI) is employed. Additionally, mid-trend fluctuations are managed through the Bollinger Bands' Width Condition.

The hypothesis proposes a trading strategy for Bitcoin based on the combination of three technical indicators: Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), and Bollinger Bands. This strategy aims to capitalise on potential trend reversals, overbought or oversold conditions, and volatility within the Bitcoin market.The entire strategy is represented as a flowchart in Figure 1. For better visibility, the flowchart is also available at: Link

2.2 Components of the Strategy

- **MACD (Moving Average Convergence Divergence):**

  Type: Momentum Indicator

  - MACD = (8-period EMA('CLOSE')) - (16-period EMA('CLOSE'))
  - Signal Line = 4-period EMA('MACD')

- **RSI (Relative Strength Index):**

  Type: Momentum Indicator

  - RSI measures the speed and magnitude of a security's recent price changes to evaluate overvalued(RSI = 80) or undervalued(RSI = 20) conditions in the price of that security.
  - In a highly trend-based BTC/USDT market the stock has a high probability of reaching overbought or oversold conditions and closing our trades in those conditions can be beneficial. By trading only when RSI is in the middle range (35 to 70) gives us an assurance of gains.

- **Bollinger Bands:**

  Type: Volatility Indicator

$$RSI_{\text{step one}} = 100 - \left[\frac{100}{1 + \frac{\text{Average gain}}{\text{Average loss}}} \right]$$

  - If there is little price action, the BBW squeezes.
  - Bollinger Band Width condition (BBW condition): The BBW condition is satisfied if the fast SMA (period 10) of the BBW is larger than the slow SMA (period 50) of the BBW. BBW condition = SMA(BBW, per=10) $>$ SMA(BBW, per=65)

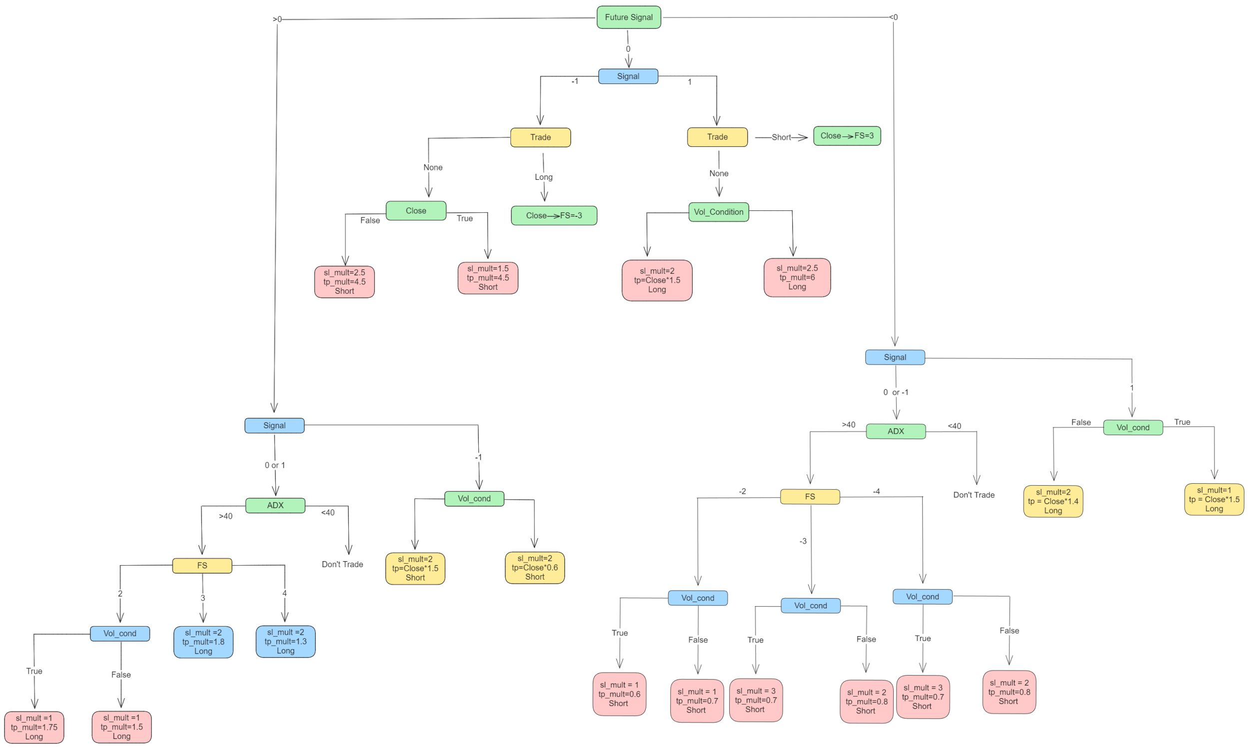


Figure 1: **Flow of the strategy** (This does not include all the entry and exit conditions. Some entry and exit conditions are also based on stoploss, takeprofit and RSI.)

$$\text{BBW} = \frac{\text{upper band} - \text{lower band}}{\text{middle band}}$$

- **Exponential Moving Average: Type:** Support line or Confirmation
  - Using a 200-period EMA, gives an assurance of the current price movement.
  - If price is above the EMA(period = 200) line then it can give confirmation for a buy signal and vice-versa.
- **Volume Condition:**

  Type: Confirmation
  - volcond = (10-period SMA('Close')) $>$ (50-period SMA('Close'))
- **Average True Range**

  Type: Price Estimator for statistically calculating the next day's highest and lowest values as well as assigning stop loss and take profit to trades. This is used to compare stop loss and take profit with the estimated price, in order to take wins before the profit level falls.

$$\text{TR} = \text{Max}\left[(\text{H} - \text{L}), |\text{H} - \text{C}_p|, |\text{L} - \text{C}_p|\right]$$

where:

$\text{H} = $ Today's high

$\text{L} = $ Today's low

$\text{C}_p = $ Yesterday's closing price

$\text{Max} = $ Highest value of the three terms

so that:

$(\text{H} - \text{L}) = $ Today's high minus the low

$|\text{H} - \text{C}_p| = $ Absolute value of today's high minus yesterday's closing price

$|\text{L} - \text{C}_p| = $ Absolute value of today's low minus yesterday's closing price

$$\frac{\text{Previous ATR}(n-1) + \text{TR}}{n}$$

where:

$n = $ Number of periods

$\text{TR} = $ True range

If there is not a previous ATR calculated, you must use:

$$\left(\frac{1}{n}\right) \sum_{i}^{n} \text{TR}_i$$

where:

$\text{TR}_i = $ Particular true range, such as first day's TR, then second, then third

$n = $ Number of periods

2.3 Signals

- **Buy Signal:**
  - **MACD crosses above the signal line below the baseline** - Suggests a trend reversal
  - **Close price $>$ 200-period EMA** - Confirms the trend direction
  - **RSI in the range 35 to 70** - Prevents investments in overbought or oversold markets that usually have a high volatility
  - **BBW condition is false [BBW condition = SMA(BBW, per=10) $>$ SMA(BBW, per=65)]** - This makes sure that the volatility of the market is low and it is not consolidating.
- **Sell Signal:**
  - **MACD crosses below the signal line above the baseline** - Suggests a trend reversal
  - **Close price $<$ 200-period EMA** - Confirms the trend direction
  - **RSI in the range 35 to 70** -Prevents investments in overbought or oversold markets that usually have a high volatility
  - **BBW condition is false** - This makes sure that the volatility of the market is low and it is not consolidating.
- **Exit Signal:**
  - **For long trades:**
    - RSI $<$ 20 (It is considered a bad indication for a long trade when the price goes into oversold condition)
    - Stop loss or Take profit levels are hit(Higher stop loss and take profit conditions have been taken if the Volume condition is satisfied)
  - **For short trades:**
    - RSI $>$ 70 (It is considered a bad indication for a short trade when the price goes into overbought condition)
    - Stop loss or Take profit levels are hit(Higher stop loss and take profit conditions have been taken if the Volume condition is not satisfied)

2.4 Further Improvement and Risk Management

- Considering the fact that the BTC/USDT market shows high volatility, there is a good chance that an exit condition shows signs for immediate trend reversal to a considerable extent.

- Hence, after closing each trade, we can improve the scope of profit by using the Average Direction Index (ADX > 40) as confirmation to trade in the opposite direction, i.e., opening a short trade after closing a long trade and vice-versa.
- Based on the conditions that closed the last trade and some other factors we start a new trade with different levels of stop-loss and take-profit.
- **Note:** These reversal based trades are started at the next candle only if we don't receive an opposing signal from the main entry conditions

2.5 Performance

- Initial Balance = 1,00,000
- Commission = 0.1 percent
- The results can be seen in table 1.
- The comparison between logged values of portfolio and close have been shown in Figure 2

Metric	Value
Return [%]	10068.655103
Return (Ann.) [%]	**214.33091**
Sharpe Ratio	1.274289
Sortino Ratio	7.321155
Calmar Ratio	8.575679
Max. Drawdown [%]	**-24.992878**
Avg. Drawdown [%]	-2.568364
# Trades	287
Win Rate [%]	64.45993
Best Trade [%]	48.480858
Worst Trade [%]	-10.756674
Profit Factor	2.524239

Table 1: Performance Metrics

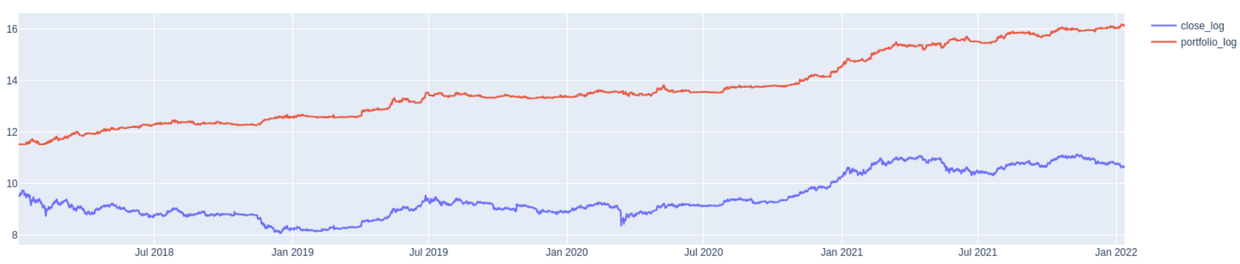
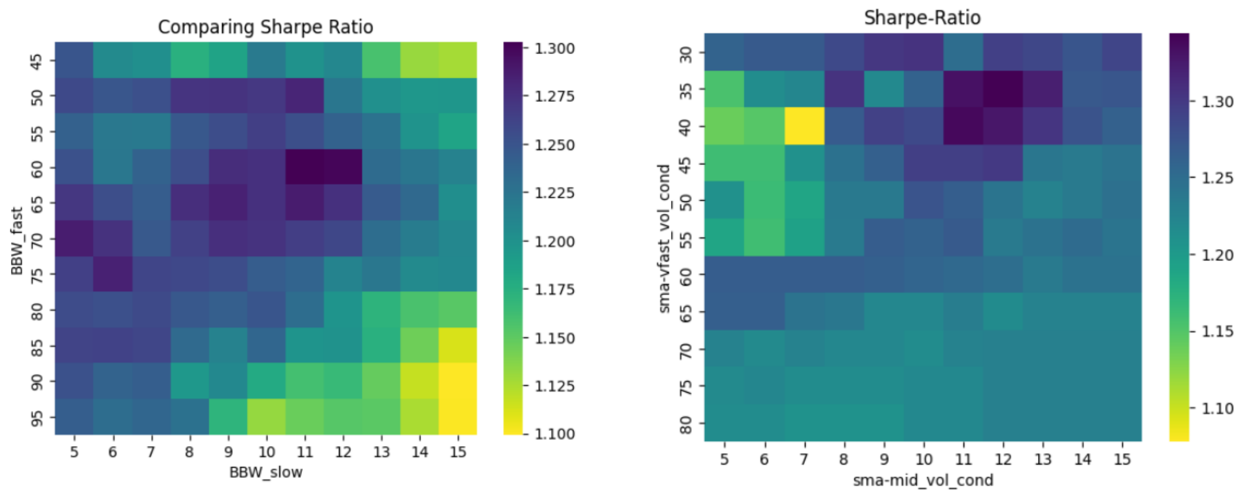


Figure 2: Logged values of portfolio as compared to logged values of close

2.6 Robustness and Consistency

- A robust strategy should not depend too much on the parameters used for the indicators. Otherwise, the chosen parameters might give a good result for the back-tested data, but not for future data since market conditions can change.

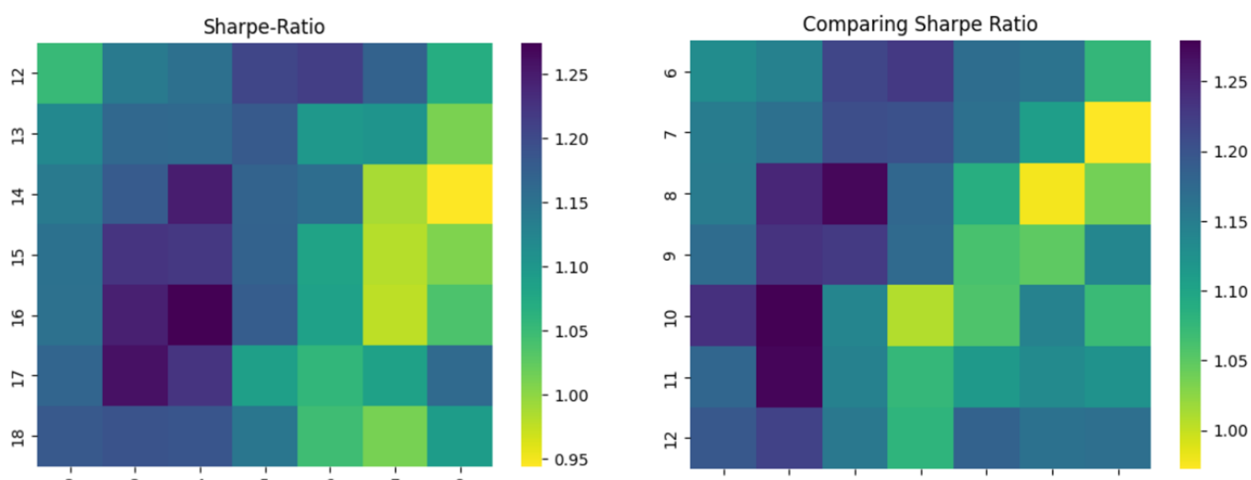
- In order to verify if the strategy was not too reliable on a particular set of parameters, a heat-map of the Sharpe ratios at different parameters was created by changing the parameter of MACD, Volume Condition and Bollinger Bands Width condition. This can be seen in the figure 3, 4 and 5.
- These heat-maps show that the sharpe ratios are good when parameters being considered are in a certain range (by value or ratio). Even if we take values different from these ranges, the sharpe ratios do not go below 0.95.
- To show the consistency of the strategy, year-by-year results of the strategy were calculated. It can be seen in Table 2



(a) Bollinger Bands Parameters

(b) Volume Condition Parameters

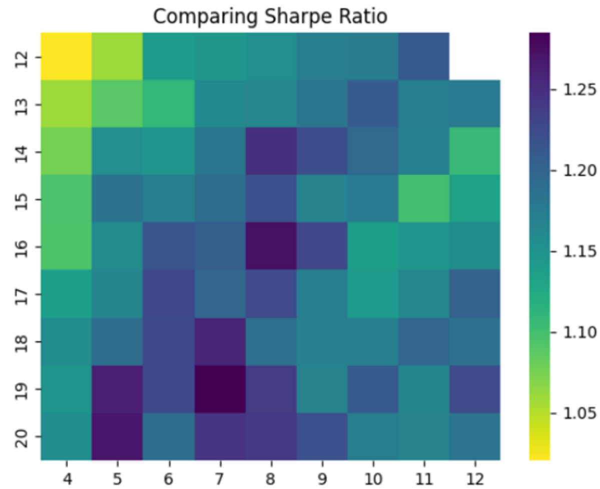
Figure 3



(a) MACD Parameters(x-axis:Signal-line period, y-axis:EMA-slow period)

(b) MACD Paramters(x-axis:Signal-line period, y-axis:EMA-Fast period)

Figure 4



(a) MACD Parameters(x-axis:EMA-fast period, y-axis:EMA-slow period)

Figure 5

Metric	First Year	Second Year	Third Year	Fourth Year
Return [%]	193.95	103.25	324.16	216.25
Return (Ann.) [%]	189.70	101.70	319.19	212.33
Sharpe Ratio	1.26	1.04	1.51	1.37
Sortino Ratio	6.41	3.89	12.24	7.61
Calmar Ratio	9.76	4.48	13.51	9.97
Max. Drawdown [%]	-19.43	-22.70	-23.62	-21.30
Avg. Drawdown [%]	-2.70	-2.83	-1.99	-2.35
# Trades	88	72	62	65
Win Rate [%]	68.18	58.33	62.90	67.69
Best Trade [%]	21.33	48.48	45.08	46.35
Worst Trade [%]	-10.36	-10.40	-10.76	-9.78
Profit Factor	2.33	2.06	3.20	2.66

Table 2: Selected Performance Metrics for Each Year

3 Further Improvements

If we are allowed to selectively start trades at the same candle we closed the previous trade. We open a new long or short trade selectively when we close a trade. The stop-loss and take-profit levels however depend on the condition that closes the last trade. The performance metrics have been mentioned in Table 3 and 4.

Metric	Value
Return [%]	163177.717792
Return (Ann.) [%]	**525.378271**
Sharpe Ratio	1.069592
Sortino Ratio	13.637528
Calmar Ratio	15.932199
Max. Drawdown [%]	**-32.975879**
Avg. Drawdown [%]	-3.066611
# Trades	232
Win Rate [%]	47.413793
Best Trade [%]	70.999689
Worst Trade [%]	-9.606236
Profit Factor	3.946149

Table 3: Selected Performance Metrics (Aggregated)

Metric	Year 1	Year 2	Year 3	Year 4
Return [%]	1362.73	251.59	202.44	547.85
Return (Ann.) [%]	1310.65	246.84	199.72	534.86
Sharpe Ratio	1.103	1.082	1.056	1.107
Sortino Ratio	34.02	7.23	5.71	14.77
Calmar Ratio	43.76	9.73	6.23	18.11
Max. Drawdown [%]	-29.95	-25.36	-32.05	-29.54
Avg. Drawdown [%]	-2.93	-3.06	-3.43	-2.53
# Trades	74	51	52	61
Win Rate [%]	59.46	33.33	36.54	55.74
Best Trade [%]	33.41	55.17	70.99	38.10
Worst Trade [%]	-6.31	-6.28	-9.61	-9.14
Profit Factor	5.96	3.13	2.67	3.96

Table 4: Selected Performance Metrics for Each Year

4 ML Based Strategies we explored

The Alternative strategies which we used employed machine learning methods, which are as follows-

4.1 EMA-SMA Crossover Strategy

- **Strategy employed**: The strategy uses a classification model to classify whether the SMA (window = 10) is greater than the EMA (window = 10) or not. Based on the change in current value and predicted value, we indicate a possible crossover of SMA and EMA. When EMA crosses over SMA, it acts as a signal to open a long trade or close an existing short trade. Further when SMA crosses over EMA, it acts as a signal to open a short trade or close an existing long trade.
- **Machine Learning Predictions**: Our approach involved forecasting the condition (SMA $>$ EMA) for the subsequent day using uni-variate time series classification models. These models categorize the time series patterns of a variable into distinct classes. Training the models involved utilizing the time series patterns derived from the differences (sma-ema) over the preceding k timestamps. This training aimed to predict whether SMA would exceed EMA on the following day.

  Among the models employed, the MUSE model from the sktime library demonstrated the highest accuracy, achieving approximately 80% Additionally, the XGBoost classifier yielded the best overall results, with an accuracy of 83% on the test data.

- **Drawback:** Using this separately does not lead to a robust strategy that has consistent results year-by-year. Although it is unconventional to test on the train data, the results drop significantly on doing so, hence we decided not to pursue this approach any further. With a stronger hypothesis that safegaurds the prediction from directly generating trade signals, we can improve the strategy further.

4.2 Forecasting Future prices:

Initially, our approach involved utilizing state-of-the-art time series prediction models to forecast future **open** and **close** prices. We used transformers and statistical models to make such predictions-

4.2.1 Transformer Forecasting

We decided to use transformers for time series forecasting, considering its efficiency in capturing long-range dependencies and complex patterns in the time series data. The architecture consists of the transformer Encoder layer, containing Layer Normalization layer, MultiHeadAttention mechanism helping the model to understand complex relations between data. Dropout layers and residual connections were added for regularization.

Since the transformer model is designed to handle raw input values effectively, normalizing the data was not necessary. For data preparation, the target column was extracted from the data frame, input and output sequences were made. We used Adam optimiser and 'mean squared error' as the loss function . We experimented with various hyper parameters such as number of transformer blocks, length of input and output sequences etc. The performance metrics on 1hr data, 30 min data and 15 min data are shown in Tables 5, 6, 7.

Metrics	**Close**	**SMA(window=10)**	**EMA(window=10)**
MAE	352	181	240
MAPE	0.007497	0.00383	0.00507

Table 5: Performance Metrics on 1 hour data

Metrics	**Close**	**SMA(window=10)**	**EMA(window=10)**
MAE	197	361	254
MAPE	0.004216	0.0075	0.005293

Table 6: Performance Metrics on 30 min data

Metrics	**Close**	**SMA(window=10)**	**EMA(window=10)**
MAE	386	160	145
MAPE	0.00803	0.003	0.003

Table 7: Performance Metrics on 15 min data

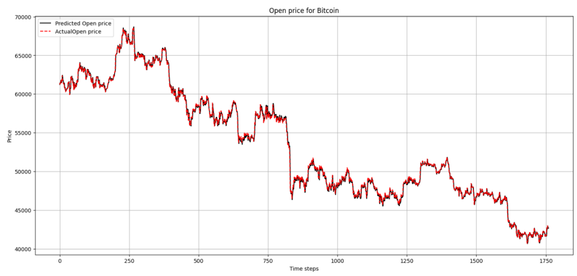
4.2.2 Statistical Based forecasting- HMM Forecasting

We used the Hidden Markov Model for forecasting the high,low,open and close prices on hourly data.We used this probabilistic statistical model to predict Bitcoin prices because it was good at finding hidden patterns and adjusting to the unpredictable nature of the cryptocurrency market.
Specifically chosen for its knack for understanding different market conditions like bullish or bearish trends, we found HMM to be great at capturing the complex movements of Bitcoin prices. For training 95% of the hourly data was used and the remaining 5% was used for testing the model.The optimal number of states was 4 and the window size was kept to be 50. Rolling predictions were made on the test data to obtain the results.

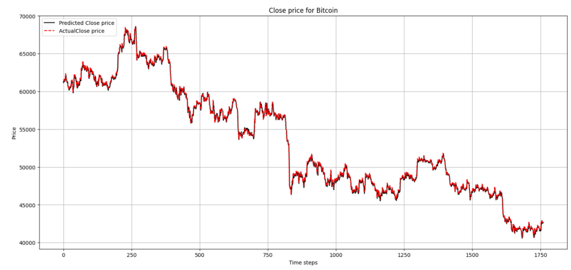
Through these forecasted results we analyzed that even though we could predict open/close prices with relatively low errors, using these results in combination with strategies proved to be ineffective,as bitcoin is a highly volatile market

Metrics	Close	Open	High	Low
MAE	264.3273481	262.63689381	242.38667235	287.08072686
MAPE	0.00499013	0.00495595	0.00456998	0.00555184
MPE	-0.00159869	-0.00065777	-0.00134443	-0.0022037
RMSE	1.26021347	1.26807613	1.11548068	1.98106811

Table 8: Performance Metrics on 1 hour Data



(a) Open Price Prediction



(b) Close Price Prediction

Figure 6: HMM Model Forecasting

and any inaccuracy in predictions could be very costly.Hence ,we opted for data smoothing. Our rationale rested on the concept that, rather than directly predicting open/close prices, we could enhance predictive capabilities by focusing on forecasting short-moving averages, specifically in terms of Short SMA and EMA. This strategic shift aimed to mitigate the impact of price volatility and improve the model's performance in capturing underlying trends.

4.3 SMA and EMA Prediction

4.3.1 FB-Prophet Model

We used the FBProphet model because :

- Prophet is robust to missing data and outliers in time-series datasets. Its ability to automatically detect and incorporate various types of seasonality in the data, including daily, weekly, and yearly patterns.
- We have maintained a train-data: test-data split of 4:1 and scaled the data between 0 and 1 before training to maintain uniformity.

Metrics	Close	EMA (Window=10)	SMA (Window=10)
MAE	344.76266	347.8038	339.5678
MAPE	56.82%	57.06%	55.90%

Table 9: Performance Metrics on 15 min data

Metrics	Close	EMA (Window=10)	SMA (Window=10)
MAE	204.0637	206.49907	206.32527
MAPE	35.84%	35.80%	35.87%

Table 10: Performance Metrics on 1 hour data

Drawbacks in Statistical models:

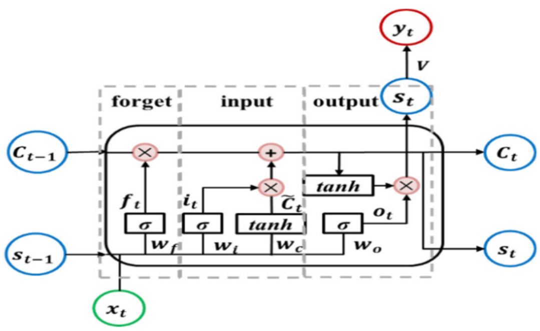
- Statistical methods make assumptions about critical data properties such as stationarity, linearity, and normality that cannot be met in practice.

- Statistical methods require a large amount of data to produce accurate results, which can be problematic if the data is limited or of poor quality.

4.3.2 Deep Learning Models

We built a simple DL model using GRU layers. The reason behind using GRU was

- GRUs are specifically designed to capture short-term dependencies in sequential data better.
- GRUs have a simpler architecture compared to long short-term memory (LSTM) networks, another type of recurrent neural network
- The gating mechanism in GRUs helps mitigate the vanishing gradient problem, which can occur in traditional RNNs during training.
- GRUs are generally easier to train than LSTMs. Their simpler architecture and fewer parameters often lead to faster convergence during training.
- GRUs often perform well on tasks with limited training data.



(a) Gated Recurrent Unit

```
Layer (type)                 Output Shape              Param #
=================================================================
gru (GRU)                    (None, 10, 50)            7950

gru_1 (GRU)                  (None, 10, 50)            15300

gru_2 (GRU)                  (None, 50)                15300

dense (Dense)                (None, 1)                 51

=================================================================
Total params: 38,601
Trainable params: 38,601
Non-trainable params: 0
```

(b) Model Summary

Figure 7: GRU - Deep Learning Model

We have maintained a train-data: test-data split of 4:1 and scaled the data between 0 and 1 before training to maintain uniformity.

Results: The results shown below are after performing inverse transform of scalar.

Metrics	Close	EMA (Window=10)	SMA (Window=10)
MAE	0.0028248	0.0012540	28.761055
MAPE	0.44%	0.18%	0.06%

Table 11: Performance Metrics on 15 min data

Metrics	Close	EMA (Window=10)	SMA (Window=10)
MAE	361.4116	140.48845	485.51133
MAPE	0.79%	0.31%	0.99%

Table 12: Performance Metrics on 1 hour data

Clearly the Deep learning model is performed better here, so we moved ahead with it.

Out of all the frequencies, the model worked well on the 15-minute data. So we used the predictions of this model and combined it with our strategy.

Metric	Value
Avg. Annual Returns (%)	**567.05645675**
Equity Final ($)	718804.37864
Equity Peak ($)	867720.21992
Return (%)	618.804379
Buy & Hold Return (%)	47.846698
Max. Drawdown (%)	**-24.544057**
Avg. Drawdown (%)	-2.144469
# Trades	128.0
Win Rate (%)	50.78125
Best Trade (%)	24.339281
Worst Trade (%)	-10.966206
Avg. Trade (%)	1.651949

Table 13: Performance Metrics

4.3.3 Hypothesis of the Strategy :

In this proposed trading strategy, the initial step involves employing a deep learning model to forecast the exponential moving average (EMA) with a window size of 10 for the upcoming trading day. Subsequently, the following conditions dictate trading decisions:

- Initiate a buy order if today's simple moving average (SMA) with a window size of 14 is greater than today's EMA with a window size of 10 and the quadratic interpolation of tomorrow's SMA (window=14) is below the predicted EMA.
- Execute a sell order if today's SMA (window=14) is less than today's EMA (window=10) and the quadratic interpolation of tomorrow's SMA (window=14) exceeds the predicted EMA.
- Any trade is initiated only when the value of the ADX indicator is greater than 25.

Train data: 2018-01-01 to 2020-12-31 (3 years)

Test data: 2021-01-01 to 2022-01-12 (1year)

The results were back tested on the test data only.

Backtesting Results:

Initial Money = 100000 **Commission** = 0.1%.

4.3.4 Drawbacks:

The predictive capabilities of the model lack the requisite accuracy for dependable trading decisions. The strategy exhibits shortcomings in effectively navigating periods of market stability, resorting to a simplistic "buy and hold" approach during such phases.

4.3.5 Future considerations:

- **Model Diversification:** Investigate and assess a range of models to identify the most suitable fit for optimizing predictive accuracy and trading performance.
- **Risk Management Implementation:** Integrate robust risk management practices to mitigate potential losses and enhance overall portfolio stability.
- **Refined Trading Conditions:** Augment the strategy with additional conditions and parameters aimed at discerning and avoiding unprofitable trades, particularly during periods of market stability. This refinement seeks to improve the strategy's adaptability to varying market conditions.

5 Results, Discussion and Conclusion

During the course of 2 months, we tried several approaches, both ML based and Price Action/Indicator based. We realized that while using ML based approaches give good results on the test data, it delivers inconsistent robustness and generalization properties. This fact led us towards Price Action/Indicator based strategies. The final strategy that we used showed great performance with good robustness and consistent results. Further improvements to the same topic can be done using either Indicator-based or ML-based conditions.