Algorithmic Trading Model for BTC/USDT Market

1. Introduction:

- A brief overview of the problem statement and the significance of developing algorithmic trading models for the BTC/USDT cryptocurrency market.
- Emphasis on the use of machine learning, statistical modeling, and programming skills to unlock the potential of ML-based algorithmic trading.

2. Problem Description:

- Explanation of the tasks involved, including data acquisition, preprocessing, model design, backtesting, risk management, and optimization.
- Highlighting the specific focus on BTC/USDT market dynamics.

3. Data and Resources:

- Mention of the availability of historical data for the BTC/USDT trading pair from January 1, 2018, to January 31, 2022.
- Encouragement for participants to use publicly available cryptocurrency market data sources, API services, or simulated data.

4. Methodology

i. Time Series Analysis:

1.

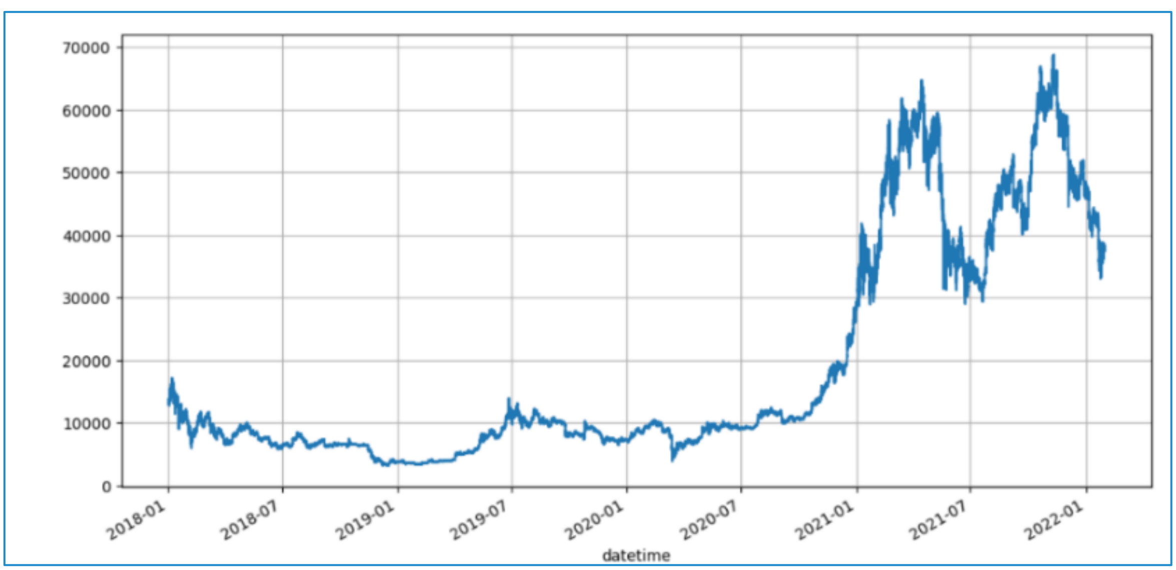


Figure 1: The time series plot from January 1, 2018, to January 31, 2022, unveils BTC/USDT closing prices, offering insights into cryptocurrency trends and dynamics.

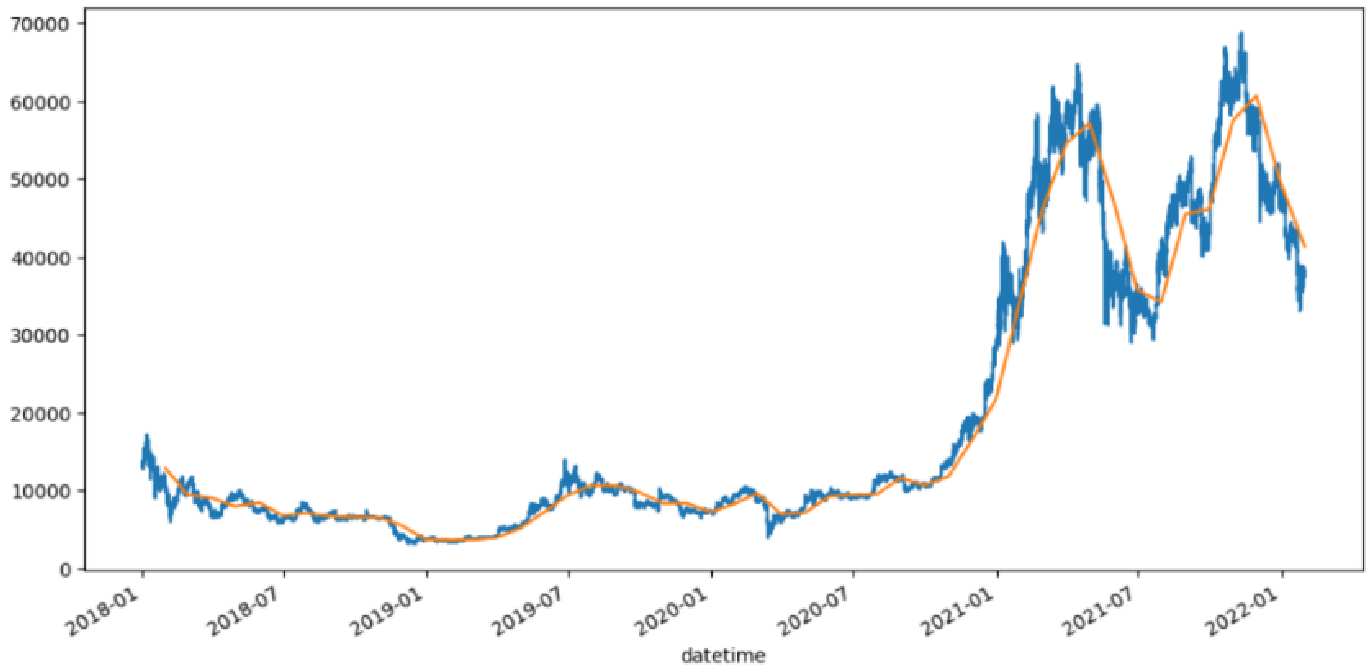


Figure 2: The plot illustrates both the daily closing prices and the monthly mean of BTC/USDT, providing insights into short-term fluctuations and long-term trends.

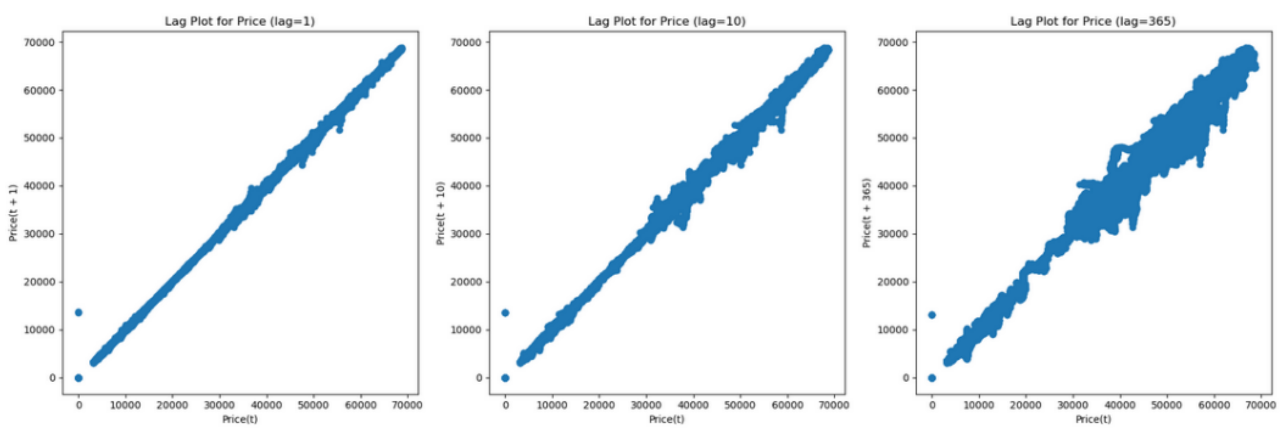


Figure 3: Lag plots depict the correlation between a time series and its lagged values, demonstrating how patterns evolve over different time intervals. As the lag increases, the correlation decreases, revealing a diminishing short-term influence on future prices.

ii. Preprocessing:

Stationarity Assessment using KPSS Test: The KPSS test results suggest that the time series is likely non-stationary. The test statistic value of 24.11 exceeds the critical values at 1%, 2.5%, 5%, and 10% significance levels, indicating a rejection of the null hypothesis of stationarity.

Details:

- Test Statistic: 24.11
- Value: 0.01
- Number of Lags: 466
- Critical Values: {'10%': 0.119, '5%': 0.146, '2.5%': 0.176, '1%': 0.216}

Interpretation: The low p-value and the significant test statistic suggest that the series is likely non-stationary. To address this, further preprocessing steps such as differencing or transformations may be necessary to achieve stationarity before modeling.

In the preprocessing stage, the time series data underwent a two-step transformation to enhance stationarity. Firstly, a logarithmic transformation was applied to mitigate the impact of varying volatility and stabilize the variance within the BTC/USDT closing prices. Following this, differencing was employed to eliminate any remaining trends and ensure the series became stationary. These sequential transformations aimed to create a more stable and predictable dataset, providing a foundation for accurate modeling of the cryptocurrency market dynamics.

Figure 4: The plot displays the log transformed BTC/USDT closing prices, aiming to stabilize the variance. The stationary test results will help assess the effectiveness of the log transformation in achieving stationarity.

Figure 5: The plot displays the log transformed BTC/USDT closing prices, aiming to stabilize the variance. The stationary test results will help assess the effectiveness of the log transformation in achieving stationarity.

iii. Models:

a. Time Series Models:

- **Moving Averages:** Utilizing simple moving averages (SMA) and exponential moving averages (EMA) to identify trends and potential entry/exit points in the BTC/USDT market.

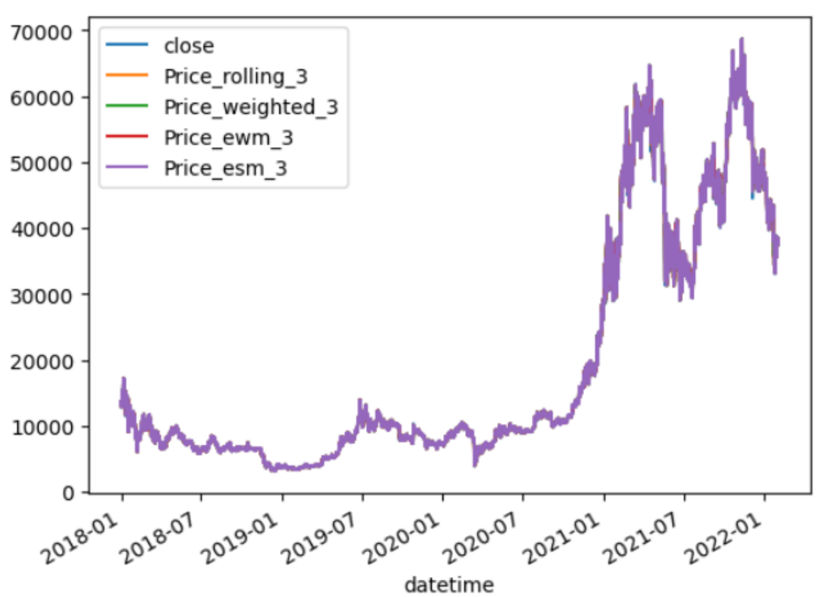
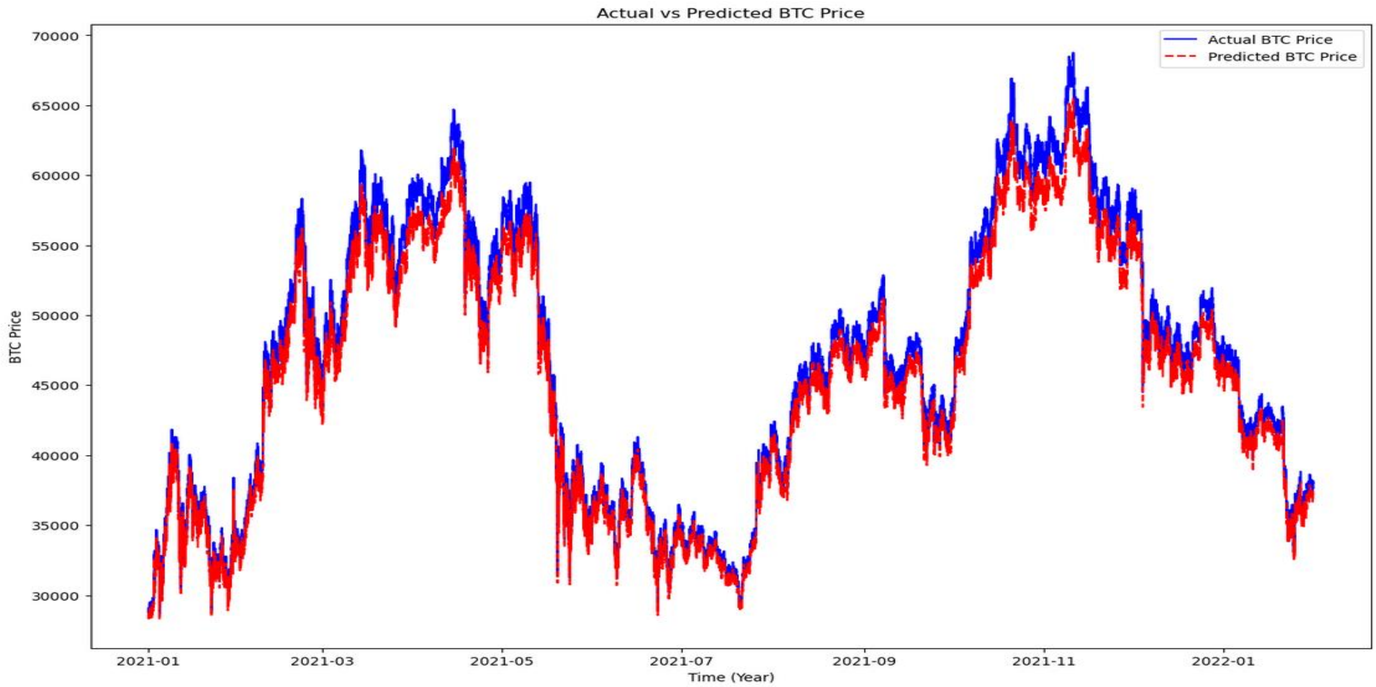


Figure 6: The graph displays BTC/USDT closing prices (blue) alongside a rolling average (orange), Weighted Moving Average (green), Exponential Weighted Moving Average (red), and Exponentially Smoothed Moving Average (purple). This concise visualization aids trend analysis and decision-making for traders.

b. Machine Learning Models:

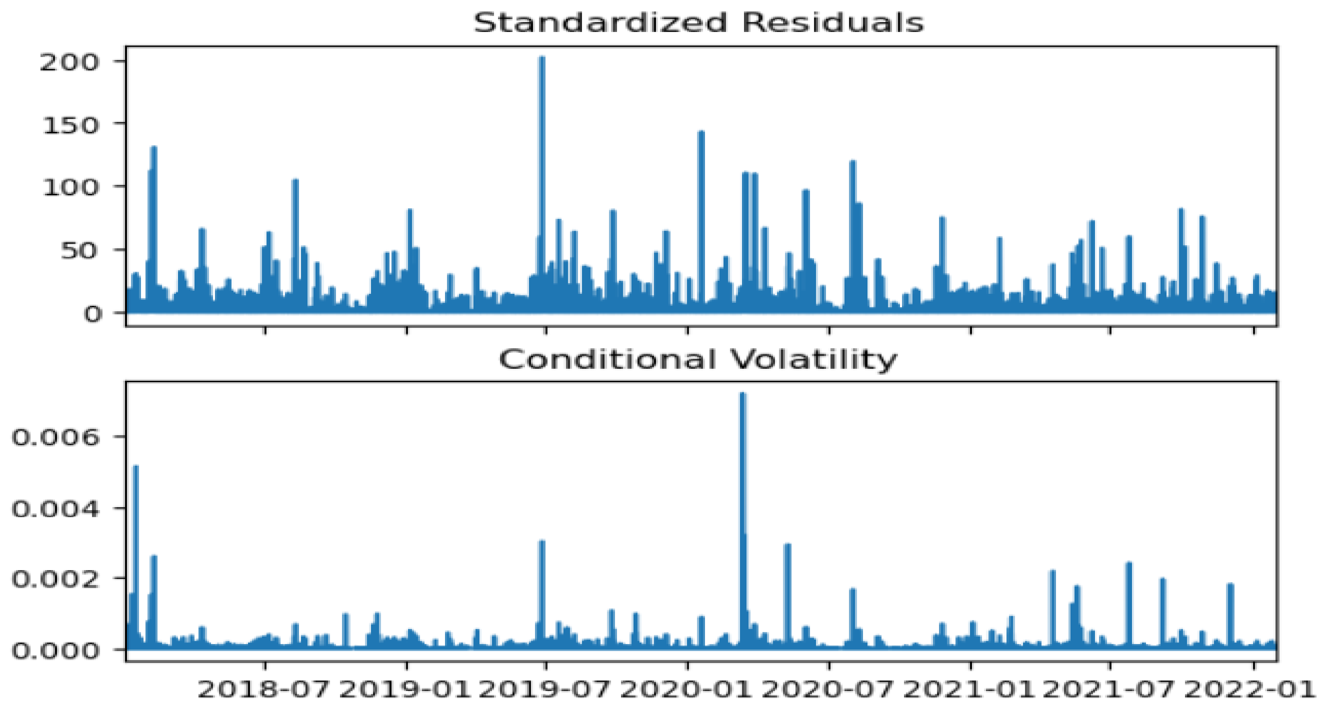
- **LSTM (Long Short-Term Memory):** Leveraging LSTM networks to capture long-term dependencies and intricate patterns in the historical price and trading volume data, enhancing the model's ability to adapt to complex market dynamics.



The LSTM model, implemented using Keras, consists of two LSTM layers, each with 50 units and a ReLU activation function. The first LSTM layer is configured to return sequences, while the second layer focuses on capturing intricate patterns. A dense layer with one unit is added for output. The model is compiled with the Adam optimizer and mean squared error loss function. It is then fitted to the test data (X_test) to assess its predictive performance. The Mean Squared Error (MSE) for the test data is computed as 2185032.805, and the R-squared (R2) score is 0.9765, indicating a high degree of accuracy and explanatory power in capturing the variance in the target variable on unseen data.

c. Risk Management:

- **GARCH (Generalized Autoregressive Conditional Heteroskedasticity):** Implementing GARCH models to assess and manage volatility, providing insights into potential risk levels associated with BTC/USDT market conditions. This includes adjusting position sizes and exposure based on predicted volatility.



The GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model was employed to capture volatility patterns in BTC squared returns. The mean model indicates a constant mean with an estimated coefficient (mu) of 4.3408e-06, suggesting a minimal average return. However, the R-squared value for the mean model is 0.000, indicating that the model does not explain much of the variation in the mean.

The volatility model, focusing on the conditional variance of the series, comprises three main parameters: omega (constant term), alpha (weight on past squared returns for volatility), and beta (weight on past conditional variances for volatility). The results show that omega is 4.1158e-11, alpha (lag 1) is 0.2, and beta (lag 1) is 0.78. These coefficients contribute to estimating the volatility over time.

The likelihood function's Log-Likelihood value is 6.91726e+06, and the model's goodness of fit is assessed using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), with values of -1.38345e+07 for both.

It's important to note that the warning about unsuccessful convergence suggests potential issues with optimization, and further investigation may be needed to ensure the model's reliability. The robust covariance estimator was employed, considering potential heteroskedasticity in the residuals.

5. Performance of Models

- Explanation of the judging criteria, including performance, model logic, risk management, code quality data visualization, presentation, and report.

6. Conclusion:

- Expression of gratitude for participating in the competition and wishing participants the best of luck in developing innovative and profitable models for the BTC/USDT market.

7. Reference

[Bitcoin Return Volatility Forecasting: A Comparative Study between GARCH and RNN]

8. Summary

MAX DRAWDOWN OF THE MODEL	**10%**
SHARPE RATIO OF THE MODEL	**0.0017**
NET PROFIT EXCEEDING BENCHMARK RETURN OF THE MODEL	**0%**
RISK-REWARD RATIO OF THE MODEL	**1.01**
MAX DURATION TIME OF SINGLE TRADE OF THE MODEL	**0.116 days**