Ethereum Price Analysis with Time Series Applications

Melih Can Kanmaz
Middle East Technical University
Ankara, Türkiye
can.kanmaz@metu.edu.tr

***Abstract*—This study aims to forecast the future prices of a cryptocurrency Ether (ETH) using the historical prices from 2016 to 2024. Implemented visualization tools to understand the data. Utilized data cleaning techniques such as anomaly detection and removal. Applied necessary transformation as Box-Cox transformation before suggesting a time series model. Checked the stationarity with KPSS, ADF and HEGY tests. To satisfy the assumption, took one regular difference. Verified the assumptions with Q-Q plots, Jarque-Bera, and ARCH tests. Fitted ARIMA, TBATS, neural networks, and exponential smoothing methods (simple, Holt's, and Holt-Winter's) to understand which performs best on this data. Evaluated the models with accuracy, MAPE, MAE and RMSE metrics.**

Keywords—ether, ethereum, price, forecast, cryptocurrency

I. Introduction

Cryptocurrencies become popular in the recent years, leading by Bitcoin and Ethereum. Similarly with all other assets people want to predict its future prices as precise as possible. While this is not a new concept for finance, the cryptocurrency market has some unique dynamics. Unlike traditional markets such as NYSE or NASDAQ, cryptocurrencies can be traded 24/7. There is no off market in whole year. In addition while still you can use you don't need a middleman to trade your asset. These two main differences redefines the rules in finance applications. Cryptocurrency market is more volatile than traditional markets. The opportunities is created and lost at blink, creating a more challenging environment for investors and traders. The volatility can be seen as variance on the historical data, creating a non-stationary series. In this study eliminating the differences in this new market, aimed to find best model for forecasting Ether (ETH) prices. The study utilized data preprocessing, stationary testing and model testing. This paper give insights on forecast of cryptocurrencies in general, contributing not only finance but also academia.

II. Related Work

Similar to the popularity of cryptocurrencies in finance, there is a popularity of cryptocurrencies in research. From popular models like ARIMA to advanced models like neural networks many studies can be found in this field. Each model has a strength on these applications, TBATS can handle complex seasonality while neural networks are better at capturing nonlinear relationships. One suggestion can be instead of looking for specific cryptocurrencies, the studies about bitcoin prices can be a better starting point since it has broader applications and is still similar to all other cryptocurrencies.

III. Methodology

A. Data Preparation

The preprocessing of the data included NA removal, outlier detection, anomaly removal and transformation. Utilized popular common Python libraries in the field as Pandas, Numpy, matplotlib, Sklearn and Scipy. Use of popular libraries enables this study to be easily replicated and open to further analysis for interested researchers.

B. Statistical Tests and Models

Stationarity is one of the most important element on a time series analysis. Before suggesting a model, testing the data with necessary tests are crucial. In this study stationary checked with KPSS, ADF and HEGY tests. The results were similar across different tests. After reaching stationary series, from the insight got from ACF and PACF plots 3 ARIMA models suggested. The related library for this tests and the models is statsmodels in Python. For further applications it can be used.

C. Forecasting

In this section the study included ARIMA, TBATS, Prophet, Neural Network and exponential smoothing models for forecasting. These models is implemented from statsmodels, forecast and tensorflow libraries in Python.

D. Evaluation Metrics

The models tested by different metrics in the study. AIC and BIC values was used for ARIMA model selection. Mean absolute percentage error, mean absolute error and root mean square error is evaluated for different models to decide on their performance. Evaluation metrics is implemented from sklearn library in Python

IV. Results

A. Stationarity Tests

Study used KPSS test and ADF test and PP (Phillips-Perron) test to check stationarity in the series. According to the p values series was not stationary as original. To determine the steps for making the series stationary we utilized the same test for trend type. While KPSS showed deterministic trend ADF and PP tests indicated stochastic trend. For seasonality detection the study used HEGY and Canova Hansen tests. Both showed there was no seasonal root in the series meaning seasonality did not exist. In addition, HEGY test results indicated a regular unit root which shows the series was not stationary. Overlapping results on stationary guided the study to take one regular difference to make the data stationary. Eventually, the p

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©2025 IEEE

values indicated there were no unit roots and one differenced series is stationary.

Test for Stationarity				
Test	Statistic	Critical Value (5%)	p-value	Result
KPSS	0,897753641	0.463	0.01	Non-stationary
ADF	-1,7221056883953200	-2,90092495	0.4197604533038781	Non-stationary
PP	-1.821	-2,9	0.370	Non-stationary

Figure 1. Tests for stationarity

Test for Trend				
Test	Statistic	Critical Value (5%)	p-value	Result
KPSS	0.0726249413781519	0.146	0.1	Deterministic Trend
ADF	-1,7463622530	-3,47074262	0.73002948551119	Stochastic Trend
PP	-2.157	-3,47	0.515	Stochastic Trend

Figure 2. Tests for trend

HEGY test results			
Statistic	Value	p-value	Significance
t1	-1,49970	0.4397	Not significant
t2	-4,16660	0	Significant
F3:4	9,57310	0.0002	Significant
F5:6	4,11660	0.0175	Not significant
F7:8	4,27060	0.0151	Not significant
F9:10	4,56370	0.0115	Not significant
F11:12	11,91760	0	Significant
F2:12	106,51210	0	Significant
F1:12	97,84280	0	Significant

Figure 3. HEGY test

Test for Stationarity of Differenced Series				
Test	Statistic	Critical Value (5%)	p-value	Result
KPSS	0.1471910641587757	0.463	0.1	Stationary
ADF	-8,08816734	-2,90147011	1.4014518544256535e-12	Stationary
PP	-8,28459881	-2,90147011	4.4254799397893563e-13	Stationary

Figure 4. Test for stationarity of differenced series

B. Model Selection

In the study ACF and PACF plots used for determining the best ARIMA/SARIMA model fits to the data. According to the plots the best ARIMA fit was ARIMA(0,1,3). To evaluate its performance and looking for the best model, study picked common time series model for comparison. This models included TBATS, Prophet, Neural Networks and exponential smoothing models. The mentioned exponential

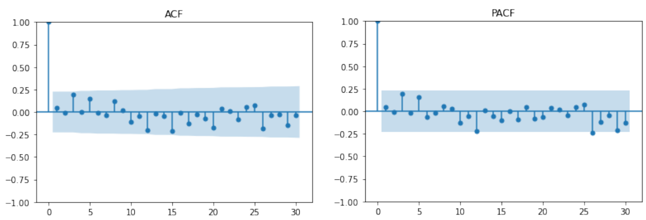

Figure 5. ACF and PACF plot for time series

Test Metrics of Exponential Smoothing Models			
Model	MAPE (%)	MAE	RMSE
SES	3,97	0,59	0.69
Holt	3,27	0,47	0.59
Holt-Winter	6,25	0.90	1,06

Figure 6. Test metrics of exponential smoothing models

smoothing models are simple exponential smoothing, Holt's exponential smoothing and Holt-Winter's exponential smoothing. To decide the best exponential smoothing study evaluated the error metrics MAPE, MAE and RMSE. The results showed that Holt's exponential smoothing has the lowest errors, therefore best exponential smoothing model for the data.

C. Forecasting Results

The first plot shows the performance of different exponential smoothing models on test data. According to their values the Holt's exponential smoothing is the best fit among exponential smoothing models. Forecasting results are evaluated by accuracy measurements such as MAPE, MAE, RMSE. According to the values of the models the best model for the data is Neural Network model.

Figure 7. Forecasts of exponential smoothing models

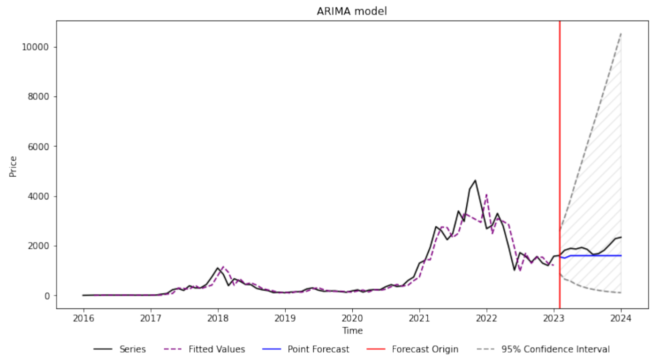

Figure 8. ARIMA model

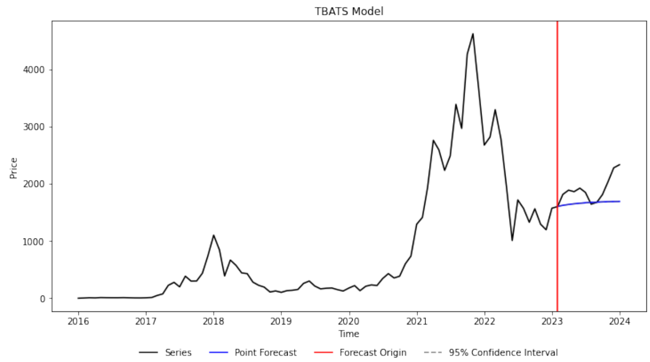

Figure 9.TBATS model

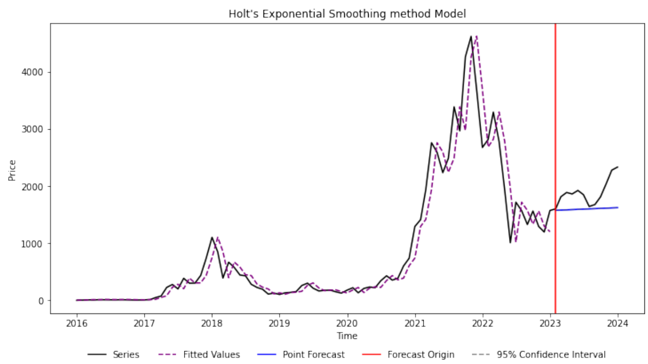

Figure 10. Holt's Exponential Smoothing model

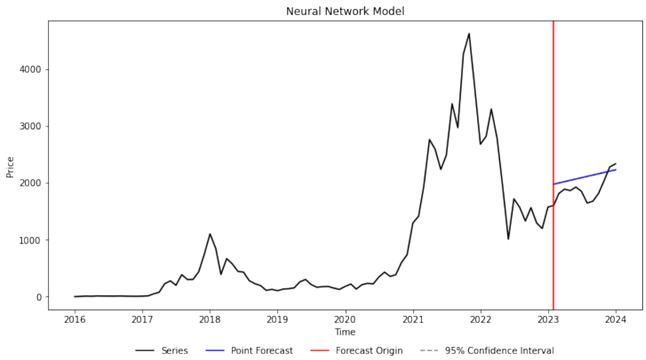

Figure 11. Neural Network model

Evalutaion Metrics for Models			
Model	**MAPE (%)**	**MAE**	**RMSE**
ARIMA	15.40	311.29	377.02
TBATS	11,54	237.43	309.59
Holt	14.53	294.91	361.77

Figure 12. Model Evaluation

V. Results

ARIMA(0,1,3) model was the best ARIMA/SARIMA fit for the model. Across the exponential smoothing methods, the Holt's exponential smoothing performed best on forecasts. The evaluation metrics MAPE, MAE and RMSE used for finding the best model for the data. According to the values, neural network provided forecasts with the least error with relatively good accuracy than other on the test data. It indicates that neural networks captured complex patterns in the data than other models tested.

VI. Discussion

Cryptocurrencies are mostly volatile assets, ETH historical price data is not different then others. The volatility creates a challenging environment for popular time series models. Existence of non-linear relationships one of the dedicator of the performance. The overall did not give high accuracies and significantly small error values on the test data; however, still we can see that across all other models Neural Network model performed best. The reason can be it captures non-linear relationships better than other models most of the times. Since it is a machine learning algorithm, it also opens a door for numerous other questions such as "Can Machine Learning models performs better than popular time series model on historical data?". In addition, to have the better performing models new variables can be used such as trading volume and market sentiment. Further analysis could be conducted to improve the models' performance.

VII. Conclusion and Future Work

This study included multiple forecasting models on the ETH historical prices dataset. ARIMA (0,1,3) was the best model across ARIMA/SARIMA models. Holt's exponential smoothing model had the best performance among exponential smoothing models. Across all other models Neural Network model had the least error on the test data. This can show, it is advantageous on volatile markets like cryptocurrencies. The result can also indicate machine learning models can perform better on forecast of cryptocurrencies. Further analysis can be conducted on additional machine learning algorithms with new variables such as trading volume and market sentiment.

References

[1] "Bitcoin & Ethereum prices (2014-2024)," *Kaggle*, Dec. 13, 2024. https://www.kaggle.com/datasets/kapturovalexander/bitcoin-and-ethereum-prices-from-start-to-2023

[2] GeeksforGeeks, "Nonlinear time series models," *GeeksforGeeks*, Jul. 05, 2024. https://www.geeksforgeeks.org/nonlinear-time-series-models/

[3] R. Adhikari and R. K. Agrawal, "An introductory study on time series modeling and forecasting," *arXiv.org*, Feb. 26, 2013. https://arxiv.org/abs/1302.6613

[4] V. Derbentsev, N. Datsenko, O. Stepanenko, and V. Bezkorovainyi, "Forecasting cryptocurrency prices time series using machine learning approach," *SHS Web of Conferences*, vol. 65, p. 02001, Jan. 2019, doi: 10.1051/shsconf/20196502001.