

Machine Learning Internship Report

Title:

Ethereum Price Forecasting using ARIMA

Intern Name:

Muhammad Hanzala Khan (Intern ID:ARCH-2506-0633)

Intern Phone no: +92 335-5191255

Organization:

Arch Technologies

Internship Duration:

June – July 2025

Abstract

This project focuses on time series forecasting of Ethereum cryptocurrency prices using the ARIMA model. The goal was to apply fundamental machine learning techniques for financial prediction. Through data cleaning, stationarity testing, and parameter tuning, we developed and evaluated an ARIMA(1,1,1) model to predict Ethereum prices for the next 30 days. The model successfully identified short-term trends and demonstrated the core concepts of time series forecasting in practice.

1. Introduction

In the field of financial analytics, predicting asset prices is a challenging but high-impact task. Cryptocurrency markets are highly volatile, making them ideal candidates for testing forecasting algorithms. This internship project applies the ARIMA model — a well-established statistical technique — to predict Ethereum prices using historical data.

2. Dataset Overview

- **Source:** Kaggle Cryptocurrency Price History Dataset
- **Asset:** Ethereum (ETH/USD)
- **Features Used:** Date, Open, High, Low, Close, Volume
- **Period Covered:** August 2015 – July 2021

- **File:** ETH-USD.csv
-

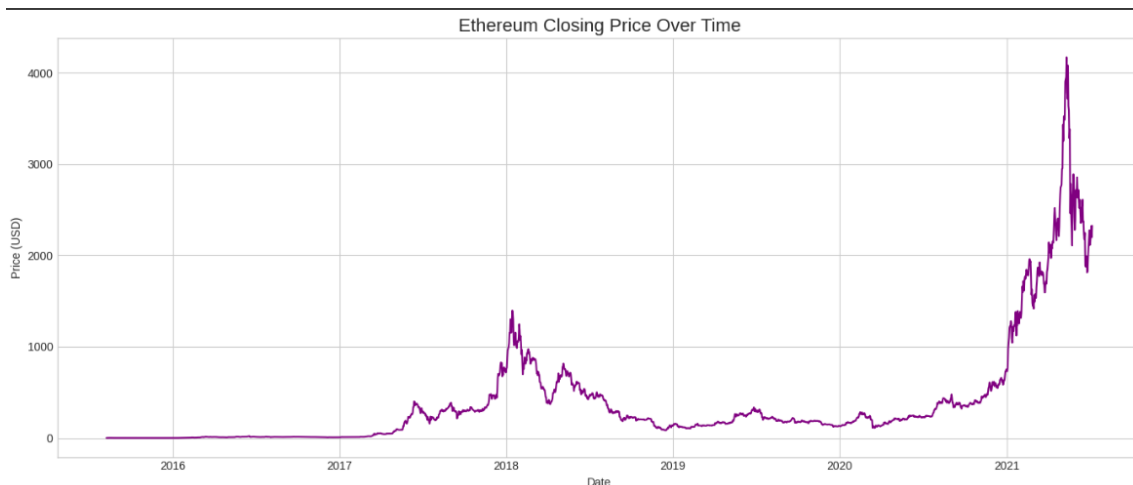
3. Methodology

Step-by-step approach:

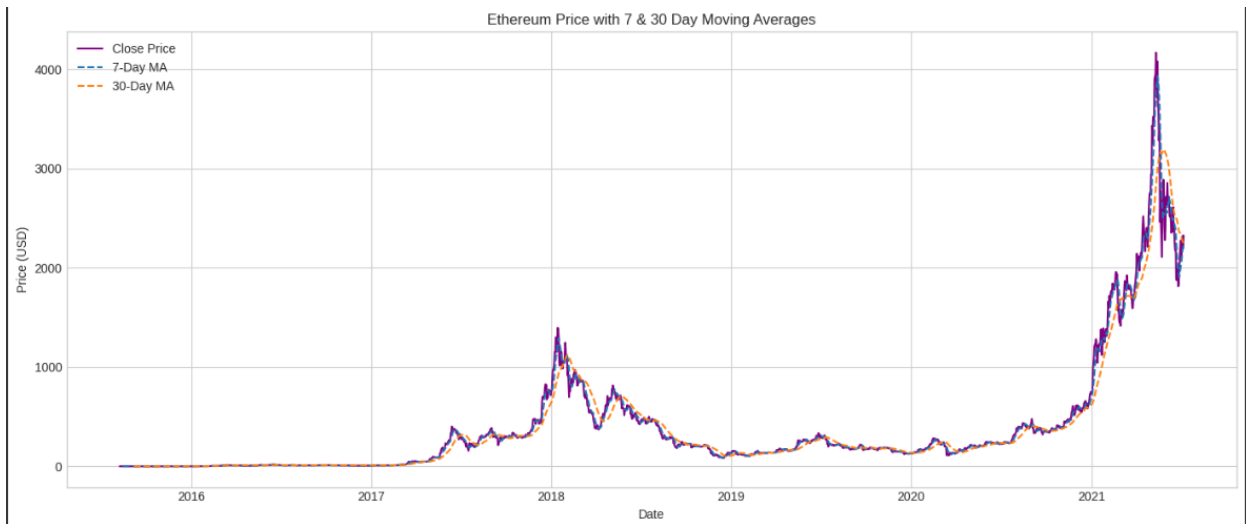
1. **Data Cleaning:**
 - Handled missing values
 - Converted Date to datetime and set as index
 2. **Exploratory Data Analysis (EDA):**
 - Plotted closing prices
 - Calculated rolling averages
 - Visualized daily percent change
 - Used correlation heatmaps
 3. **Stationarity Testing (ADF):**
 - Initial series was non-stationary (p-value ≈ 0.95)
 - Differenced the series (1st order)
 - New p-value $\approx 5.8e-17 \rightarrow$ stationary
 4. **Model Selection:**
 - Used ACF & PACF plots
 - Selected: ARIMA(1,1,1)
 5. **Model Training & Forecasting:**
 - Trained model using statsmodels
 - Forecasted Ethereum prices for 30 future days
 - Plotted predictions vs recent actual data
-

4. Results

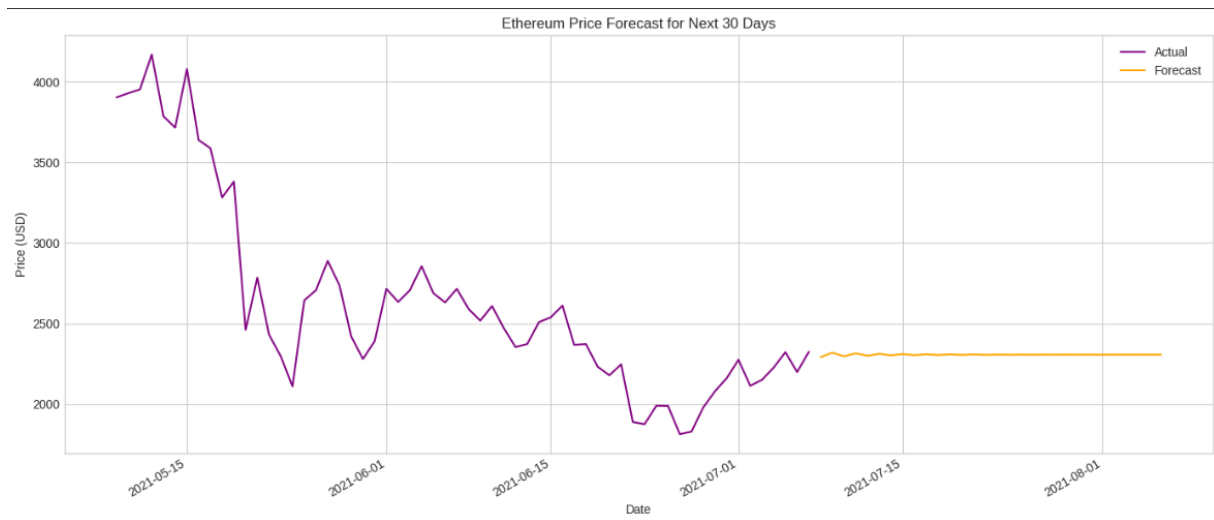
- The ARIMA(1,1,1) model successfully captured Ethereum's short-term pricing trend
- Forecast plot aligned closely with recent actual price patterns
- Charts included:
 - Time series visualization



- Rolling average comparison



- Forecasted price plot (last 60 days + 30 day forecast)



5. Conclusion

This project demonstrated the implementation of a classical machine learning algorithm on real-world financial data. The ARIMA model effectively predicted future prices over a short horizon, making it suitable for time series applications. However, cryptocurrency markets are inherently noisy — future improvements could involve:

- Seasonal ARIMA (SARIMA)
- LSTM neural networks
- Facebook Prophet for trend/seasonality modeling

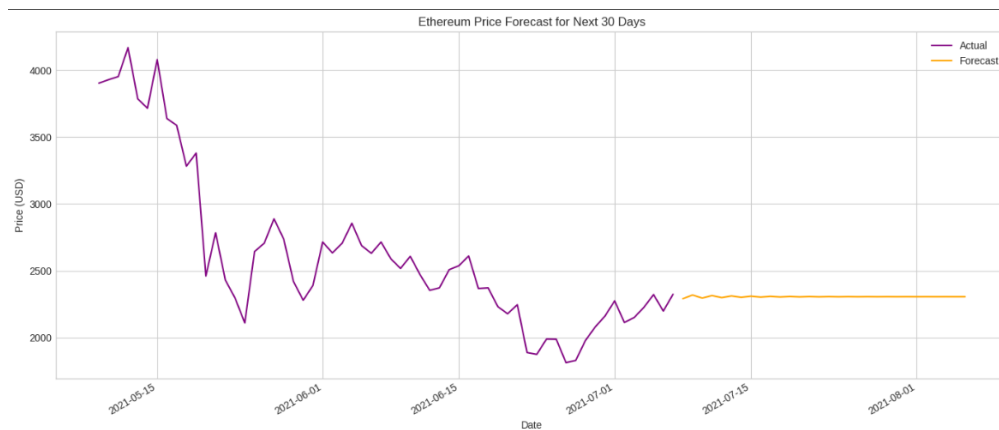
6. Tools Used

- Python 3.10 (Google Colab)
- Pandas, NumPy
- Matplotlib, Seaborn

- Statsmodels
 - Jupyter Notebook
-

Appendix

A. Forecast Chart Screenshot



B. Code Notebook

The complete .ipynb notebook includes:

- Markdown headings
- Clean, structured code
- Step-wise implementation

GitHub link: https://github.com/Hanzala095/Ethereum_ARIMA_Forecasting

LinkedIn post: [https://www.linkedin.com/posts/muhammad-hanzala-khan-4081492b7_machinelearning-timeseries-arima-activity-7339586838325571585-](https://www.linkedin.com/posts/muhammad-hanzala-khan-4081492b7_machinelearning-timeseries-arima-activity-7339586838325571585-dPNv?utm_source=social_share_send&utm_medium=member_desktop_web&rcm=ACoAAEwEfwgBawGQQzJsiLaIdlILkdo5conJl_c)

[dPNv?utm_source=social_share_send&utm_medium=member_desktop_web&rcm=ACoAAEwEfwgBawGQQzJsiLaIdlILkdo5conJl_c](https://www.linkedin.com/posts/muhammad-hanzala-khan-4081492b7_machinelearning-timeseries-arima-activity-7339586838325571585-dPNv?utm_source=social_share_send&utm_medium=member_desktop_web&rcm=ACoAAEwEfwgBawGQQzJsiLaIdlILkdo5conJl_c)
