POC- Etherum price prediction using on-chain analysis

Table des matières

Abstract

Here we're implementing a predictive price strategy that relies only on on-chain signals created by Ethereum «*whales*». The goal is to show that even with a data sample of just one year, it is possible to forecast significant price changes. Our assumption is that monitoring the transfer of large quantities of tokens from a whale address to exchanges can signal impending sell-offs.

1. Introduction

1. 1 Whales monitoring

Cryptocurrencies like Ethereum are volatile assets influenced significantly by large stakeholders, known as "whales", typically defined as wallets holding substantial balances, sufficient to significantly impact market liquidity and akin to insiders in traditional financial services.
In Ethereum, a handful of addresses control over half of the total supply, giving these entities unparalleled power in shaping market dynamics.

1.2 On Chain analysis

As all transactions and holdings are recorded on public ledgers, on-chain data provides a unique opportunity to analyze their behavior in real-time. We will use services such as blockchain explorers and RPC nodes read developments on the ledger to effectively monitor all incoming and past transactions as new blocks and transactions constantly update the decentralized ledgers of cryptocurrencies. As on-chain signals are always publicly available, we can make an analysis of insider trading on cryptocurrencies.

1.3 Price movement anticipation

Unlike traditional methods that rely on technical indicators or historical price patterns, the aim here is to delves into blockchain-derived data to construct a predictive framework. By focusing on Ethereum whales, we try to answer the question whether the activities of these large holders be used to anticipate price movements effectively by using only the transparent nature of blockchain technology that enables granular tracking of transactions, balances, and interactions, providing invaluable insights into the activities of influential market participants

2. Approach

2.1 Methodology

Our analysis focuses on the Ethereum blockchain, leveraging API like Etherscan and Alchemy for data collection. Whales are identified as non-project wallets within the top 5% of active Ethereum holders. The dataset includes daily data from approximatively 50 randomly selected whale wallets in 2022.

2.2 Signals tracking

Two signals are analyzed:

- *Ethereum balance change*s: Historical balances reconstructed by aggregating transaction data, as direct balance queries were not available. Significant increases or decreases in balance were flagged as potential indicators of whale activity.
- *Ethereum-to-stablecoin ratios:* such as USDT, USDC, and DAI were tracked to calculate the ratio of Ethereum to stablecoins held by each wallet. This ratio serves as a proxy for market sentiment, with higher stablecoin holdings often indicating a bearish outlook

2.3 LSTM for price prediction based on those signals

We then build a machine learning model that predicts price changes based on whale activities. A *Long Short-Term Memory (LSTM)* neural network was employed for price prediction, leveraging its ability to model temporal dependencies in volatile markets and to operate well with noisy and incomplete data.

2.4 Evaluation of the model

To assess the efficacy of the whale-based model, it was compared against a momentum trading strategy. The momentum model relied solely on historical price data, providing a benchmark for evaluating the added value of on-chain signals.

3. Implementation

We collected whale addresses, their signals, and finally the design of our price prediction model.

3.1 Whale wallets address retrieving

By our definition of whale wallets, we searched for

- Active addresses with an Ethereum balance in the top 5%.
- Through Alchemy's RPC endpoints:
  - we randomly selected 10 blocks mined during 2022,
    - collected their lists of transactions, then
      - used those transactions to create a database of unique active wallets.
        - We then picked the top 5% of the richest wallets to designate as our whales, resulting in approximately 50 unique whale addresses directly from the blockchain

3.2 Signal collection

Using the list of whales wallets, we
collect daily data signals
from each address about their
Ethereum balances and
Ethereum to stablecoin holding ratio during 2022.

3.2.1 Historical Ethereum Balances

We collect historical Ethereum balance of a wallet using methods provided by web3.eth Python library and Etherscan's API which returned an address's balance at a given block that allowed us to collect the transaction history of each of our whales.
Following our asumption, shifting Ethereum balances of whales can aprovide insights into market manipulation.

3.2.2 Ethereum to Stablecoin Ratio

Focusing on the three most popular stablecoins: USDC, USDT, and DAI, we retrieve data about our whale's stablecoins by using Moralis dedicated endpoint to track ERC 20s. For each whale, we recorded the amount of stablecoins they owned each day. Using our previously collected data on Ethereum balances, we are able to create a new signal, following our assumption that the stability of stablecoins makes them attractive to traders who want to avoid the risk of price volatility and protect their investments against price fluctuations by converting their cryptocurrencies into stablecoins during times of market volatility.

3.3 Price Prediction Model

We begin by preprocessing our whale signals through normalizing and scaling it to ensure that the model is better able to capture temporal patterns. The use of LSTM will then help us identify patterns and trends in the data that can be used to make price forecasts.
Through the validation data set, we determine that approximately 80 epochs of training with one hidden layer appears to be optimal as our validation MSE no longer decreases as much beyond that point. Although the model is not perfect and fails to capitalize on the many ups and downs of the market, it appears adept at sensing significant signals.

4. Conclusion

Thiis notebook presents a POC for cryptocurrency price prediction using on-chain whale data, and shown promising insights despite only having a year's worth of data to train on. By collecting and analyzing on-chain data, we were able to capture important information about whale behavior and use it to make accurate predictions about future price movements.

References

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4184367

https://docs.coinmetrics.io/

https://chatgpt.com/