**Twitter or Social Media Volume Momentum (w/ crypto?)**

**Data Sources:**

* **Twitter API (Tweepy)**: Collect tweets mentioning specific cryptocurrencies or hashtags.
* **Reddit API (Pushshift)**: Scrape mentions and comments from forums like **r/CryptoCurrency**.
* **Alternative Data Services**: Platforms like **LunarCRUSH** and **Santiment** specialize in crypto social volume data.

**Model Training & Evaluation:**

* **Feature Engineering**:
  + Tweet volume, retweets, likes, sentiment scores for tweets related to specific assets.
  + Sentiment scores using NLP models like **VADER** or **TextBlob**.
  + On-chain metrics as additional features (e.g., wallet activity).
* **Neural Network Approach**:
  + Use an **LSTM** network to identify spikes in social media volume and their relationship to price action.
  + Alternatively, use **transformers** (like **BERT**) to deeply analyze tweet sentiment to predict buy/sell signals.
* **Deployment**:
  + Deploy a **Python script** to continually scrape data and feed it into the model.
  + **Real-Time Evaluation**: Track spikes in social volume and automatically trigger trades if thresholds are reached.
* **Evaluation**:
  + **Accuracy of Buy Signals**: Calculate how many times a social volume spike leads to significant price movement.
  + **Hit/Miss Ratio**: Monitor whether a predicted price move materializes.

The way in which I imagine doing it is:

First, using the API of Dex Screener (or a similar service if you know one which is more capable) to spot the wallets which take (and hold) surprisingly large positions very early on in the life of a coin which subsequently skyrocketed (their win-rate is vast). Then, I want to transfer those to Cielo (or a similar service) to screen for the wallets which make transactions rarely. By this logic, wallets which make rare but big and accurate transactions are most likely insiders.

This would be the cornerstone of the analysis. Each time one of those wallets makes a new acquisition, I want to go onto the second element of the analysis: keyword analysis. I want to scrape posts and discussions involving that specific coin and use some sort of natural language processing algorithm to evaluate whether the conversation truly relates to my coin of choice. Then I want to use one of the pre-trained algorithms to assess the sentiment concerning that particular coin.

The investment decision would then be made based on the output of a neural network model which takes the input of the sentiment and some other relevant variables (like the returns of that coin and the crypto market in general (BTC))