Project Proposal: Ethereum Weekly Volatility Prediction Using Market, Sentiment, and On-chain Data

Project Summary

This project aims to predict Ethereum weekly volatility in the short term utilizing real-time Ethereum information from the Yahoo Finance API and Ethereum on-chain data from the Etherscan API, together with the Kaggle dataset that include sentiment indicators. Based on the information and dataset, a machine learning classification model will be generated to determine whether the following week's Ethereum's closing price will rise or fall. The model will apply preprocessing and data cleaning, feature engineering, and evaluation and visualization. This project illustrates how market sentiment and network usage dynamics work together to influence Ethereum volatility movement.

Data Sources Table

#	Name / Short Description	Source URL	Type	List of Fields	Format	Have tried to access/ collect data with python? yes/no	Estimated data size
1	Ethereum Market Data (Yahoo Finance)	https://finance.yahoo.com/quote/ETH-USD/history	API	Date, Open, High, Low, Close, Adj Close, Volume	CSV/JSON	yes	~2,000 rows
2	Crypto Fear & Greed Index	https://www.kaggle.com/datasets/liiucbs/crypto-fear-and-greed-index	File	date, fng_value, fng_classification	CSV/JSON	yes	~1,500 rows
3	Ethereum On-chain Data	https://docs.etherscan.io/introduction	API	Gas metrics, Tx volume, Rewards, Token flows	JSON	yes	~2,000 rows

Type of Analysis / Model

This project will develop a binary classification model that will predict whether Ethereum prices will increase or decrease the following week. The target variable will be taken from the daily closing prices (1 = increase, 0 = decrease). Independent features will include technical factors (such as moving averages, returns), sentiment factors from the Fear & Greed Index,

as well as Ethereum on-chain factors (such as Gas metrics, Tx volume, Rewards, Token flows). Data preprocessing will include filling missing values, normalizing, as well as feature selection. Models such as Logistic Regression and XGBoost will be trained, evaluated, and compared. We will test the models for accuracy, precision, recall, as well as ROC-AUC for evaluation purposes. The analysis will endeavor to determine which factor influences changes in Ethereum prices the most in the short term, combined with visualization to better understand how Ethereum market behavior, investor sentiment, and network usage dynamics are related.