

Final Project Report

Project Title: Cryptocurrency Liquidity Prediction for Market Stability

1. Executive Summary This project successfully developed a machine learning model to predict cryptocurrency liquidity, with the goal of providing an early warning system for market instability. Utilizing market data from CoinGecko for March 2022, a Random Forest Regressor was trained to forecast a custom liquidity ratio. The model was deployed via a local Streamlit application, demonstrating a functional end-to-end pipeline for data processing, feature engineering, and prediction.

2. Project Objective The primary objective was to leverage machine learning techniques to analyze historical market data and build a predictive model capable of detecting early signs of liquidity crises. By forecasting liquidity, the system aims to empower traders and financial institutions with insights needed to manage risk effectively.

3. Methodology

- **1. Data Preprocessing:** The dataset was cleaned to handle inconsistencies, and key date columns were converted to a usable format. Numerical features were scaled to prevent model bias.
- **2. Exploratory Data Analysis (EDA):** Visual analysis using bar plots and heatmaps revealed key market leaders and identified a moderate correlation between price and trading volume, validating the feature set.
- **3. Feature Engineering:** Two domain-specific features were created: a 3-day price moving average (price_ma) to capture short-term trends and a liquidity_ratio to serve as the prediction target.
- **4. Modeling:** A Random Forest Regressor was selected for its robustness and ability to handle non-linear relationships in the data.
- **5. Deployment:** A user-friendly web interface was built using Streamlit, allowing for interactive, real-time liquidity predictions.

4. Performance Evaluation The model's performance was assessed using standard regression metrics:

- **R² Score: 0.19.** This indicates that the model explains approximately 19% of the variance in the liquidity ratio. While modest, this is a reasonable starting point given the limited dataset (two days of data) and market volatility.
- **Root Mean Square Error (RMSE): 14.69.**
- **Mean Absolute Error (MAE): 2.25.**

These metrics suggest that while the model has predictive power, its accuracy is constrained by the small size of the training data.

5. Conclusion and Future Work The project successfully delivered a proof-of-concept for liquidity prediction. The model performs reasonably well within the limitations of the dataset and provides a valuable tool for local analysis through its interactive interface..