Final Report: Cryptocurrency Liquidity Prediction for Market Stability

Project Objective

The primary goal of this project is to predict cryptocurrency liquidity levels to enhance market stability and help stakeholders such as traders, exchanges, and institutions make informed decisions.

Liquidity, defined as the ease of buying/selling assets without significant price change, is a critical factor in financial markets. Low liquidity often leads to volatility and market crashes. This project leverages machine learning techniques to build a predictive model using historical market data.

Data Description

- Source: Google Drive dataset (20162017).
- Raw Data Includes:
  - Price data (open, high, low, close)
  - Trading volume
  - Market cap

Data Preprocessing

1. Missing Value Treatment: Forward-fill and mean imputation applied.
2. Normalization: Min-Max scaling on numerical features.
3. Encoding: Date and categorical features converted.
4. Outlier Handling: Applied Winsorization and IQR filtering.

Feature Engineering

- Rolling statistics: Moving averages (7/30-day), rolling std dev

- Liquidity Ratios: Volume/Market Cap ratio

- Volatility Measures: Daily returns, momentum

- Technical Indicators: RSI, MACD

Exploratory Data Analysis (EDA)

- Strong correlation found between volume and liquidity.

- Visualizations: heatmaps, line plots, histograms, boxplots

- Observed periodic trends aligned with external events

Modeling Approach

Models Evaluated:

- Linear Regression

- Decision Trees

- Random Forest (Selected)

- Gradient Boosting

- XGBoost

Selected Model: Random Forest Regressor

Model Evaluation (Actual Results)

- Test Mean Squared Error (MSE): 5.31

- Test R Score: 0.635

Interpretation:

- The model explains about 63.5% of the variance in liquidity.

- RMSE suggests moderate prediction errors, acceptable for financial data.

Hyperparameter Tuning

Used GridSearchCV on:

- n_estimators, max_depth, min_samples_split

Best Parameters:

- n_estimators: 150

- max_depth: 12

- min_samples_split: 5

Deployment

- Built using Streamlit

- Allows real-time prediction of liquidity based on user input

- Input: Volume, price change, volatility

- Output: Predicted liquidity score

Final Insights

- Volume and volatility are the strongest predictors of liquidity.

- The model provides practical forecasts to avoid liquidity crises.

Deliverables

- Cleaned and Feature Engineered Dataset

- Trained Random Forest Model (Saved as .pkl)

- Streamlit UI for local deployment

- Evaluation metrics JSON

- Documentation: HLD, LLD, Pipeline, Final Report

Conclusion

The model demonstrates that historical and engineered features can predict liquidity levels with reasonable accuracy. The pipeline supports real-time use through a simple interface, enabling informed trading decisions and risk mitigation.