

Cryptocurrency Liquidity Prediction - Project Report

1. High-Level Design (HLD)

This machine learning project predicts cryptocurrency liquidity ratio, a key factor for market stability, using structured market data. The system includes:

- Data collection (historical crypto metrics)
- Data preprocessing (cleanup, normalization)
- Feature engineering (liquidity-related metrics)
- Model training using Linear Regression
- Streamlit-based prediction app for local deployment

Technologies Used: Python, pandas, scikit-learn, Streamlit, joblib

2. Low-Level Design (LLD)

Input: Two CSV datasets (2022-03-16 and 2022-03-17) containing:

- Price, 1h/24h/7d % change
- Volume and market cap

Preprocessing:

- Dropped missing values
- Normalized numerical features using MinMaxScaler
- Parsed 'date' into datetime format

Feature Engineering:

- $\text{cap_to_volume} = \text{mkt_cap} / \text{volume}$
- $\text{weighted_change} = 24\text{h} * \text{volume}$
- $\text{liquidity_ratio} = \text{volume} / \text{price}$

Modeling:

- Algorithm: Linear Regression
- Trained on 8 features
- Target: liquidity_ratio

Deployment:

- Streamlit web app with sliders
- Loads trained model (.pkl)
- Predicts and displays liquidity instantly

3. Pipeline Architecture

[Raw Data (CSV)]

[Preprocessing]

- Drop missing values
- Normalize price, volume, market cap

Cryptocurrency Liquidity Prediction - Project Report

- Convert date to datetime

[Feature Engineering]

- Compute:
 $\text{cap_to_volume} = \text{mkt_cap} / \text{volume}$
 $\text{weighted_change} = 24\text{h change} \cdot \text{volume}$
 $\text{liquidity_ratio} = \text{volume} / \text{price}$

[Model Training]

- Input: 8 features
- Algorithm: Linear Regression

[Model Export]

- Save trained model using joblib liquidity_predictor.pkl

[Deployment]

- Streamlit app loads model
- User inputs market values
- App predicts liquidity ratio

4. Model Performance

Model: Linear Regression

Features used: 8 total (including engineered ones)

Target: liquidity_ratio

Metrics on training data:

- RMSE: Low
- MAE: Minimal
- R Score: High (expected overfit due to small size)

5. Exploratory Data Analysis (EDA)

- Distribution of price: Skewed left (most coins have low price)
- Liquidity vs. Volume: Positive correlation
- Heatmap: Strong correlations:
 - liquidity_ratio 24h_volume, mkt_cap

6. Deployment

Tool: Streamlit

Execution: streamlit run app.py

Cryptocurrency Liquidity Prediction - Project Report

Features:

- Input sliders for user to simulate coin market stats
- Predicts and displays liquidity
- Loads pre-trained .pkl model (no retraining)