High-Level Design (HLD) Document

1. Project Overview

The Cryptocurrency Liquidity Prediction System is designed to predict whether a cryptocurrency will face a liquidity crisis based on market indicators such as price, trading volume, and market capitalization. The system uses historical snapshots of market data, performs feature engineering to create liquidity metrics, and trains a machine learning model to classify liquidity levels.

2. Problem Statement

Cryptocurrency markets are volatile and prone to liquidity crises. Identifying low-liquidity assets early helps investors and exchanges mitigate risks. The goal of this project is to develop a predictive model that classifies cryptocurrencies into high or low liquidity categories.

3. Scope

The project covers data collection, cleaning, feature engineering, model development, and deployment. The model predicts liquidity using historical CoinGecko data and can be extended to real-time use cases.

4. System Architecture

The architecture consists of four main layers: 1. Data Layer: Handles collection and storage of cryptocurrency snapshots. 2. Processing Layer: Cleans and preprocesses data, performs feature engineering. 3. ML Layer: Trains machine learning models and evaluates performance. 4. Deployment Layer: Provides a user interface through Streamlit for predictions.

5. Technologies Used

Component	Technology
Data Collection	CoinGecko CSV snapshots
Data Processing	Python, Pandas, NumPy
Visualization	Matplotlib, Seaborn
Modeling	Scikit-learn, RandomForest
Deployment	Streamlit, pyngrok

6. Data Flow Overview

1. Input CSV files are loaded into the system. 2. Preprocessing cleans missing values and standardizes formats. 3. Features such as Volume-to-Market Cap ratio are engineered. 4. Machine learning model is trained on labeled data. 5. Model predictions are deployed through a Streamlit interface.