Low-Level Design (LLD) Document

Project Title: Cryptocurrency Liquidity Prediction for Market Stability

1. Module Breakdown:

A. Data Collection:

- Source: coin_gecko dataset (CSV)

- Method: Manual download from provided link

B. Data Preprocessing:

- Scripts: clean_data.py, data_preprocessing.py

- Tasks: Null handling, duplicate removal, data type conversion

C. Feature Engineering:

- Scripts: feature_engineering.py

- Features: Moving averages, volatility, liquidity ratios

D. Exploratory Data Analysis (EDA):

- Notebook: Eda.ipynb

- Visuals: Histograms, correlation heatmaps, trend plots

E. Model Training & Evaluation:

- Notebook: model_training_and_evaluation.ipynb

- Models: Random Forest, Linear Regression, XGBoost

- Metrics: RMSE, MAE, R

F. Hyperparameter Tuning:

- Script: hyperparameter_tuning.py

- Tool: GridSearchCV

G. Model Selection & Testing:

- Script: model_selection.py, test_model.py

H. Deployment:

- Script: deployment.py

- Framework: Flask/Streamlit

2. File Structure:

- src/
  - data_preprocessing.py
  - feature_engineering.py
  - model.py
  - deployment.py
- model/
  - best_random_forest_model.pkl
  - scaler.pkl
- notebooks/
  - EDA, feature engineering, model training

3. Data Flow:

Raw Data Cleaned Data Feature Engineered Data Model Training Evaluation Deployment