High-Level Design Document

Cryptocurrency Liquidity Prediction Project

Contents

1 Document Overview **2**

2 Project Objective **2**

3 System Architecture Overview **2**

4 Technology Stack **3**

5 Data Flow Diagram **4**

6 Security and Scalability **4**

7 Assumptions and Constraints **4**

8 Conclusion **5**

1 Document Overview

This document outlines the high-level system design for a cryptocurrency liquidity prediction project using machine learning. It covers system components, architecture, data flow, and key technologies.

2 Project Objective

To build a machine learning system that predicts the liquidity of various cryptocurrencies using historical data such as prices, volume, and temporal patterns. This helps monitor and prevent potential liquidity risks in volatile crypto markets.

3 System Architecture Overview

The overall architecture consists of the following major components:

- Raw data ingestion from CoinGecko CSVs
- Preprocessing pipeline to clean and standardize the dataset
- Feature engineering using rolling statistics and time-based features
- Model training using regression techniques (e.g., XGBoost)
- Evaluation using MAE, RMSE, and R^2 metrics
- Local deployment using Streamlit

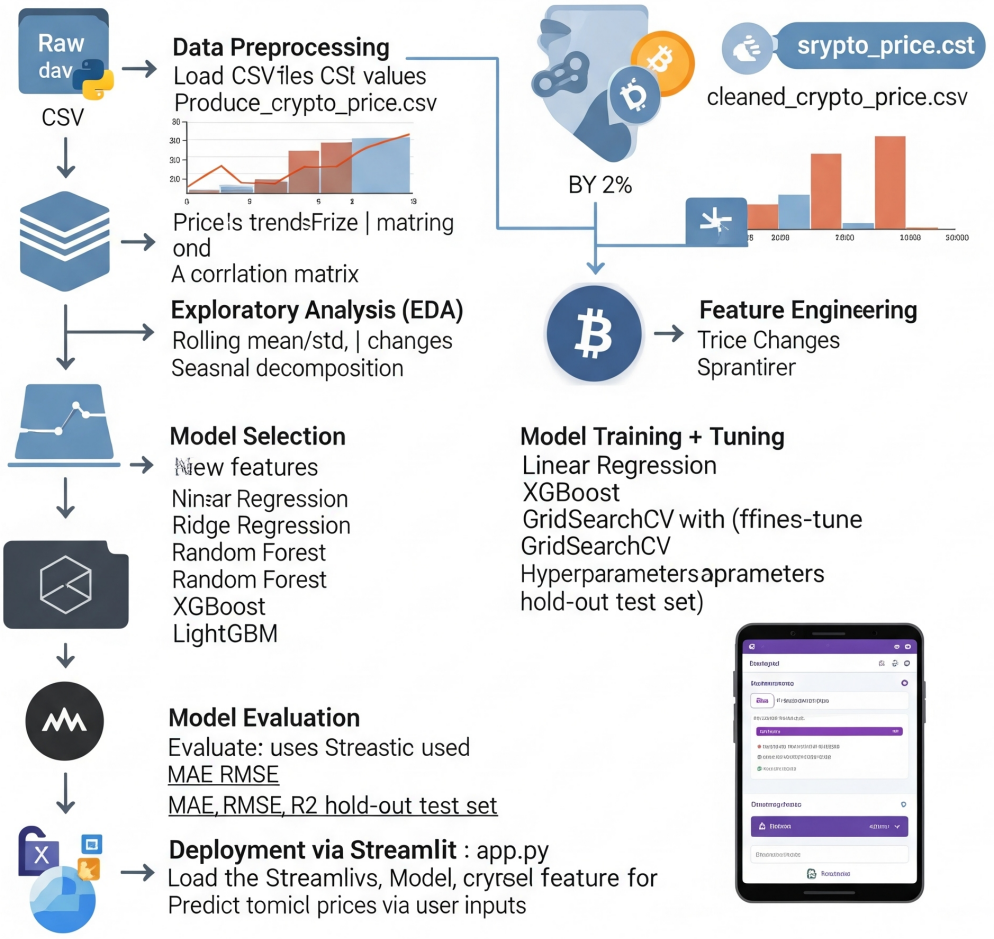

Figure: System Architecture Overview

4 Technology Stack

- **Programming Language:** Python 3.x
- **Data Handling:** Pandas, NumPy
- **Visualization:** Matplotlib, Seaborn
- **Modeling:** Scikit-learn, XGBoost, LightGBM
- **Deployment:** Streamlit (local)
- **Storage:** CSV files (local disk)
- **Version Control:** GitHub

5 Data Flow Diagram

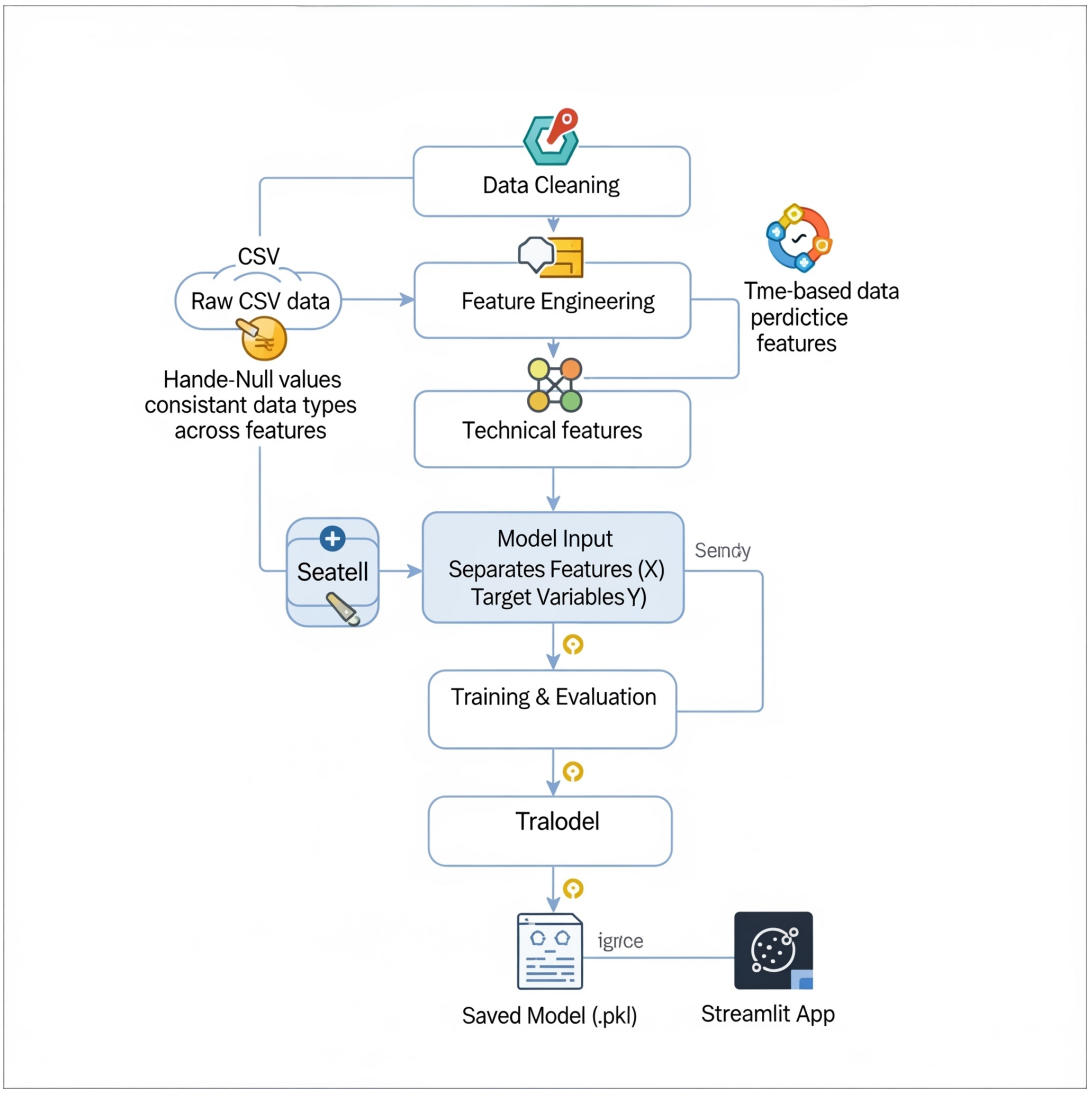

Figure: Data Flow from Raw to Prediction

6 Security and Scalability

- Local-only processing: No external API or database access
- No PII used or stored
- Easily extendable to cloud/real-time prediction pipelines
- Code modularity ensures flexibility and reusability

7 Assumptions and Constraints

- Input data is assumed to be consistently structured over time.
- Model only supports daily-level data; intraday prediction is out of scope.
- Time-based and price-based engineered features are precomputed.

- Streamlit app is intended for local testing only (no public hosting yet).

8 Conclusion

This high-level design captures the essential components and structure of the cryptocurrency liquidity prediction system. It serves as a foundational document for development, validation, and deployment.