# HIGH-LEVEL DESIGN (HLD)

## Cryptocurrency Liquidity Prediction for Market Stability

*Machine Learning Project*

Version 1.0

# 1. Executive Summary

This document presents the high-level design for a machine learning system that predicts cryptocurrency liquidity levels to support market stability analysis. The system leverages historical cryptocurrency market data (2016-2017) to classify cryptocurrencies as having either high or low liquidity, enabling traders and financial institutions to make informed decisions and manage risks effectively.

The solution employs supervised machine learning classification algorithms, with Logistic Regression achieving 95.15% accuracy on test data. The system is deployed as a Flask web application that provides real-time liquidity predictions based on user-provided market parameters.

# 2. System Overview

## 2.1 Problem Statement

Cryptocurrency markets exhibit high volatility, and liquidity plays a critical role in market stability. Insufficient liquidity leads to increased price fluctuations, making it difficult for traders to execute transactions without significantly impacting prices. This system addresses the need for early detection of liquidity crises by predicting liquidity levels based on market indicators.

## 2.2 Objectives

- Predict cryptocurrency liquidity levels (High/Low) with high accuracy
- Provide insights into market stability patterns through exploratory data analysis
- Enable traders and exchanges to make informed risk management decisions
- Deploy an accessible web interface for real-time liquidity predictions

## 2.3 Scope

**In Scope:**

- Historical cryptocurrency data from 2016-2017
- Binary classification (High Liquidity vs Low Liquidity)
- Web-based prediction interface using Flask
- Model training, evaluation, and deployment

# 3. System Architecture

## 3.1 Architecture Overview

The system follows a standard machine learning pipeline architecture with four major components: Data Processing Layer, Model Training Layer, Model Serving Layer, and Presentation Layer. The diagram below illustrates the complete system architecture:

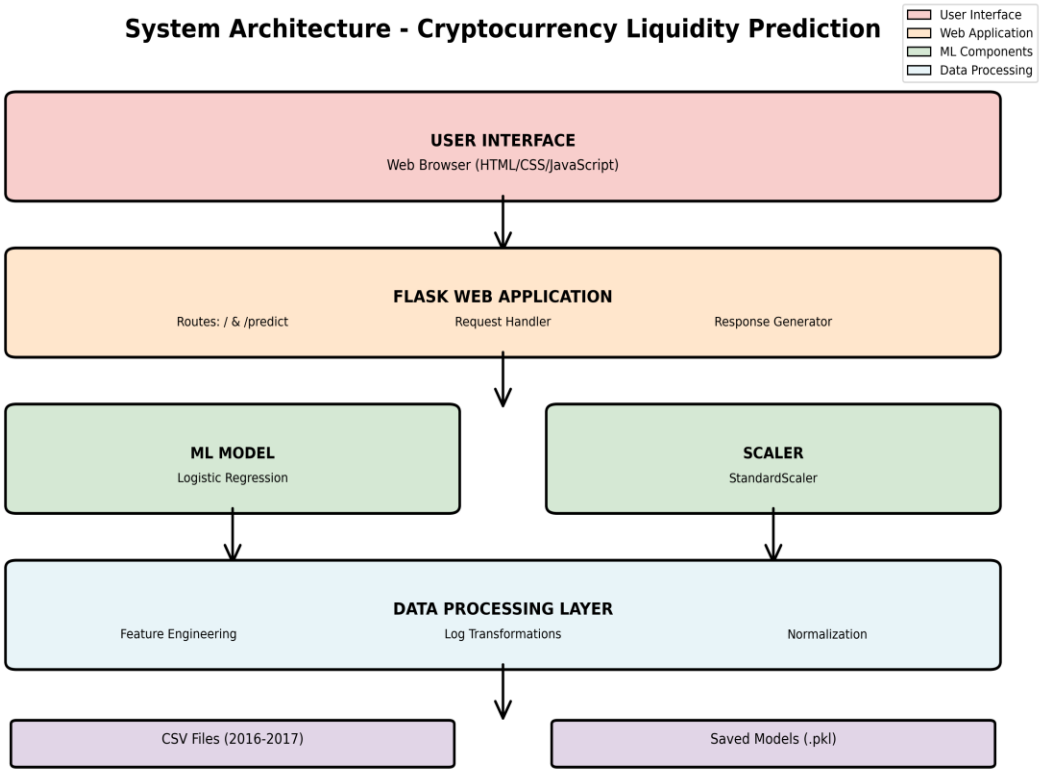


*Figure 1: System Architecture Diagram*

## 3.2 System Components

| Component | Description |
|---|---|
| Data Layer | Handles data ingestion, cleaning, and preprocessing. Merges 2016 and 2017 datasets, handles missing values, and performs feature engineering. |
| Feature Engineering | Creates derived features including logarithmic transformations (log_volume, log_mkt_cap), liquidity ratios, and day-over-day price changes. |
| ML Pipeline | Trains multiple classification models (Logistic Regression, Random Forest, Gradient Boosting, SVM), performs hyperparameter tuning, and evaluates model performance. |
| Model Storage | Serializes trained models and scalers using joblib for persistence and deployment. |

| | |
|---|---|
| Flask Web App | Provides REST API endpoints for prediction. Loads trained model and scaler, accepts user inputs (price, volume, market cap), and returns liquidity predictions with confidence scores. |
| Web Interface | HTML/CSS/JavaScript frontend that allows users to input cryptocurrency parameters and displays prediction results with visual indicators. |

## 3.3 ML Pipeline Flow

The complete machine learning pipeline from data collection to deployment is illustrated below:

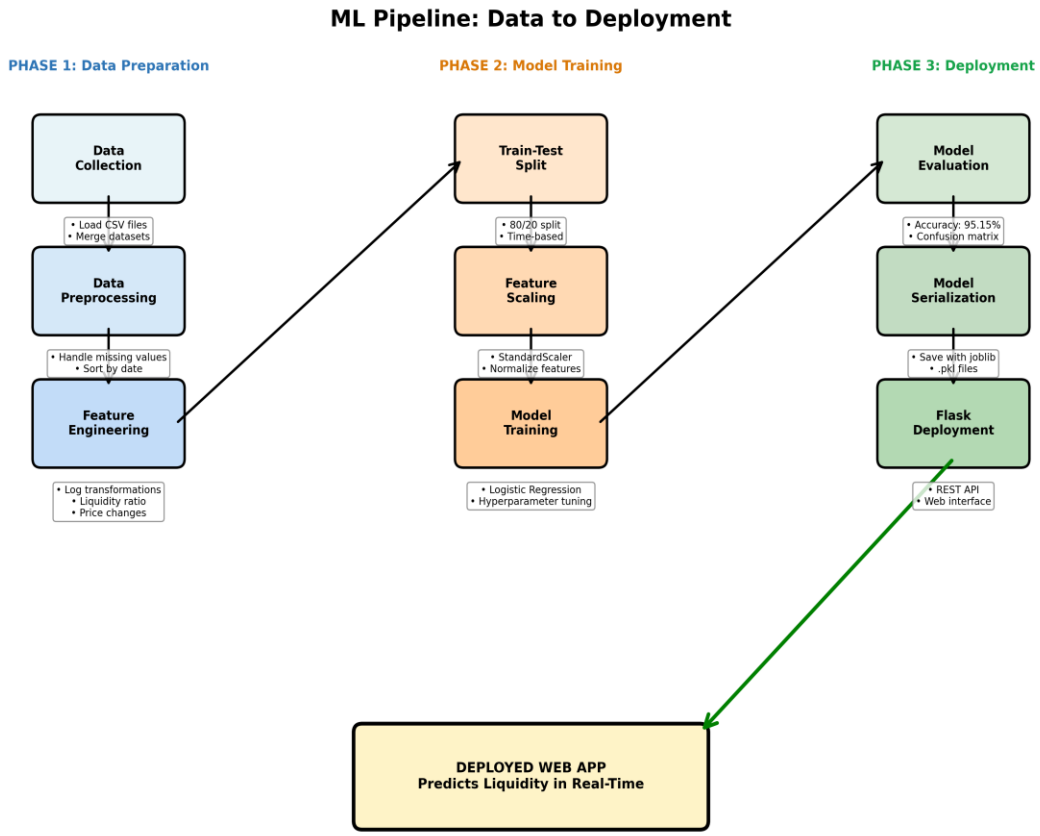


*Figure 2: Machine Learning Pipeline*

# 4. Machine Learning Approach

## 4.1 Model Comparison and Selection

Four classification algorithms were evaluated to identify the optimal model for liquidity prediction. The comparison chart below shows the accuracy achieved by each model:

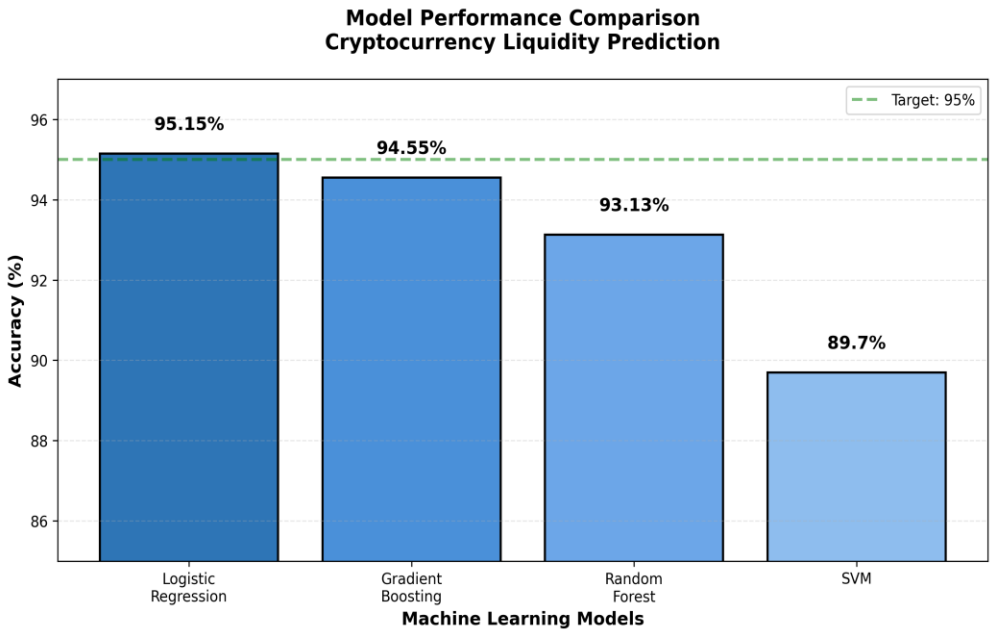


*Figure 3: Model Performance Comparison*

**Selected Model:** Logistic Regression was chosen as the final model due to its superior accuracy (95.15%), computational efficiency, and interpretability. The model provides clear decision boundaries and probability estimates for each prediction.

## 5. Conclusion

This high-level design document outlines a comprehensive machine learning system for predicting cryptocurrency liquidity levels. The system achieves 95.15% accuracy using Logistic Regression and provides an intuitive web interface for real-time predictions. The architecture is modular, maintainable, and designed for future enhancements including real-time data integration and advanced analytics.

The solution effectively addresses the problem of early liquidity crisis detection, enabling traders and financial institutions to make informed risk management decisions in the volatile cryptocurrency market.