Machine Learning Project Cryptocurrency Liquidity Prediction for Market Stability

Submitted by: Rohit Raikwar



Pipeline Architecture Document

**1. Data Ingestion**

* Load historical cryptocurrency data including price, volume, and other relevant market indicators.
* Source data from CSV files or external APIs.

**2. Preprocessing**

* Handle missing values using techniques like interpolation or imputation.
* Normalize and scale numerical features to ensure consistent data ranges.
* Perform Exploratory Data Analysis (EDA) to understand data patterns, distributions, and outliers.

**3. Feature Engineering**

* Create technical indicators such as:
  + Moving Averages (SMA, EMA)
  + Volatility indicators (standard deviation, Bollinger Bands)
  + Liquidity ratios (volume-to-price ratios, spread analysis)
* Encode categorical features and apply lag features where necessary.

**4. Model Training**

* Train multiple machine learning models:
  + RandomForestRegressor
  + XGBoostRegressor
  + Optional: LSTM for sequence-based predictions
* Use cross-validation and hyperparameter tuning to optimize models.

**5. Evaluation**

* Evaluate models using standard regression metrics:
  + Root Mean Squared Error (RMSE)
  + Mean Absolute Error (MAE)
  + R-squared (R² Score)
* Visualize prediction vs. actuals and residual plots.

**6. Deployment**

* Deploy the selected model using **Streamlit**.
* Build an interactive web application that:
  + Accepts new input data
  + Displays prediction results dynamically
  + Visualizes trends, model performance, and forecasted liquidity levels

