

Pipeline Architecture and Documentation

Pipeline Overview

- **Ingestion:**
Loading raw dataset CSV.
- **Preprocessing:**
 - Data cleaning (drop/ impute missing values).
 - Encoding and feature scaling.
- **Feature Engineering:**
Adding moving averages, volatility, ATR, Bollinger Bands.
- **Train-Test Split:**
Random or time-based split for validation.
- **Model Training:**
 - Random Forest training and feature importance extraction.
 - LSTM sequence creation and training.
- **Evaluation:**
Calculate MAE, RMSE, R^2 on test data.
- **Output:**
Final metrics report and trained models for prediction.

Pipeline Components

- Data Loader Module
- Preprocessing Module
- Feature Engineering Module
- Model Training Module
- Evaluation Module
- Prediction Module

Execution Flow Diagram

Draw a flowchart connecting above modules in the execution order.