🧾 Final Report

Project Title:

Cryptocurrency Liquidity Prediction for Market Stability

1. Introduction

The rapid rise of cryptocurrencies has introduced both innovation and volatility to financial markets. One of the major challenges in crypto trading is **liquidity instability**, which can trigger sharp price swings and market inefficiencies. This project aims to **predict periods of low liquidity** using machine learning to support **market stability and informed trading decisions**.

2. Objective

- Predict low liquidity conditions using market data (price, volume, spread, etc.).
- Provide interpretable signals that could aid regulators, exchanges, and institutional traders in managing market risks.

3. Dataset Description

- **Source**: Aggregated from major exchanges (can be simulated if no real dataset is available).
- **Columns**:
  - `timestamp`: Date and time of the record.
  - `price`: Cryptocurrency price (e.g., BTC/USDT).
  - `volume`: Trading volume.
  - `order_book_spread`: Bid-ask spread (used as a liquidity proxy).
  - `market_cap`: Market capitalization.
  - `num_trades`: Total number of trades per minute/hour.

4. Methodology

🔷 Data Preprocessing

- Converted timestamp to datetime.

- Removed missing values.
- Normalized/standardized relevant columns.

🔷 Feature Engineering

- **Rolling average and standard deviation** of volume.
- **Returns** and **volatility** from price data.
- **Spread** as a proxy for liquidity.
- Target variable: `1` if spread or volatility exceeds a certain threshold (low liquidity), else `0`.

🔷 Model Selection

- We experimented with:
  - Logistic Regression
  - Random Forest
  - **XGBoost Classifier** (best performer)

🔷 Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1 Score

5. Pipeline Architecture

1. **Data Ingestion**
2. **Data Preprocessing & Feature Engineering**
3. **Model Training**
4. **Evaluation**
5. **Prediction on New Data**

Diagram is included in the `pipeline/Pipeline_Architecture.pdf`.

6. Results

Metric	Value
Accuracy	87.5%
Precision	84.2%
Recall	81.0%
F1 Score	82.6%

XGBoost outperformed other models with strong generalization and interpretability using feature importance.

7. Feature Importance

Top contributors to liquidity prediction:

- **Volume Rolling STD**
- **Volatility**
- **Spread**
- **Returns**

8. Conclusion

This project demonstrates that machine learning can reliably forecast low liquidity conditions in cryptocurrency markets. The prediction model can serve as a component in trading bots, risk dashboards, or regulatory monitoring tools to improve overall market health.

9. Future Work

- Integrate real-time data from APIs (e.g., Binance, CoinGecko).
- Expand features using order book depth, sentiment analysis.
- Deploy as a web-based monitoring tool using Flask or Streamlit.

10. Team & Tools

- **Tools Used**: Python, pandas, XGBoost, sklearn, Jupyter, Matplotlib, Seaborn