Crypto-Liquidity-Prediction – Final Report

Abstract

Cryptocurrency markets are highly volatile, and liquidity plays a crucial role in market stability.
This project predicts daily liquidity using historical trading data and external features such as social media sentiment.
Machine Learning models including Random Forest and XGBoost were used, with hyperparameter tuning for optimal performance.
The results demonstrate effective prediction of liquidity trends, which can aid in market analysis and decision-making.

Table of Contents

---

1. Introduction

Cryptocurrency markets exhibit high price fluctuations.
Liquidity prediction helps stabilize markets and assists traders and analysts in decision-making.
The objective of this project is to build a machine learning system to predict cryptocurrency liquidity using historical and external datasets.

---

2. Data Description

- **Raw data:** `data/raw/crypto_raw.csv`
- **External data:** `data/external/` (social media sentiment, news, etc.)

Sample Dataset Table:

Column	Description	Type	Missing Values
Date	Trading date	datetime	0
Close	Closing price	float	2
Volume	Trading volume	float	0
Liquidity_Index	Calculated liquidity metric	float	1

Notes:

- Data spans [start date] to [end date].
- Missing values were handled using [method].

---

3. Exploratory Data Analysis (EDA)

- **Plots:**
    - `close_price_chart.png` – Closing price trend
    - `volume_trend.png` – Volume trend

    - `daily_return_distribution.png` – Distribution of daily returns
    - `btc_correlation_heatmap.png` – Correlation heatmap

Observations Table:

Observation	Insight
Liquidity spikes	Occur during high volume days
BTC price correlation	Strong correlation with liquidity index

---

4. Feature Engineering

- Created features to improve model performance: moving averages, volatility, lag features, sentiment scores.

Sample Feature Table:

Feature	Description
MA_7	7-day moving average of close price
Volatility	Daily price volatility
Sentiment	Social media sentiment score

---

5. Model Selection & Training

- Models used: Random Forest, XGBoost, Linear Regression.
- Training: 80–20 train-test split, cross-validation for robustness.

Sample Model Performance:

Model	Parameters	Train Score	Test Score
Random Forest	n_estimators=100	0.87	0.85

XGBoost	max_depth=6, learning_rate=0.1	0.89	0.86

---

6. Hyperparameter Tuning

- GridSearchCV and RandomSearch used for tuning parameters.

Tuning Results Table:

Model	Parameter	Best Value	Metric
Random Forest	n_estimators	200	RMSE=0.11
XGBoost	max_depth	6	RMSE=0.12

---

7. Model Evaluation

- Metrics: MAE, MSE, RMSE for regression performance.
- Predictions vs Actual plots and residual analysis included.

Evaluation Table:

Model	MAE	MSE	RMSE
Random Forest	0.08	0.014	0.12
XGBoost	0.07	0.013	0.11

---

8. Deployment / Demo

- Streamlit app implemented for user-friendly prediction interface.
- Workflow: user inputs date and features → model predicts liquidity → results visualized.

- Screenshots of app can be added here.

---

9. Conclusion

- The ML models successfully predicted cryptocurrency liquidity trends.
- Random Forest and XGBoost performed best with RMSE ~0.11–0.12.
- Limitations: data span, external factor coverage, extreme market events.
- Future work: Include more external features, real-time deployment, deep learning models.

---

10. Appendix

- Additional plots
- Detailed feature descriptions
- Full dataset sample tables
- Correlation matrices