Cryptocurrency Volatility Prediction using Machine Learning

This report presents an end-to-end machine learning system designed to predict short-term cryptocurrency volatility using historical market data.

Abstract

Cryptocurrency markets are highly volatile, making risk assessment challenging. This project predicts short-term volatility using OHLC prices, trading volume, and market capitalization. The best-performing model is deployed using Streamlit.

Problem Statement

To predict short-term cryptocurrency volatility using historical market data to support risk management and informed decision-making.

Dataset Description

The dataset contains daily historical records of multiple cryptocurrencies, including open, high, low, close prices, trading volume, and market capitalization.

Methodology

Data Preprocessing: Removal of redundant columns, date normalization, chronological sorting, and missing value handling.

Feature Engineering: Daily returns, rolling volatility, price spreads, and liquidity ratio.

Models: Linear Regression (baseline) and Random Forest Regressor (advanced).

Evaluation Metrics: RMSE, MAE, and R^2 score.

Results

Random Forest Regressor outperformed Linear Regression by achieving lower error values and a positive R^2 score, indicating better non-linear pattern learning.

Feature Importance

Price range features and liquidity indicators were found to be the most influential in predicting cryptocurrency volatility.

Deployment

The trained Random Forest model was deployed locally using Streamlit, providing an interactive web interface for volatility prediction.

Conclusion

The project successfully demonstrates a complete machine learning pipeline from data preprocessing to deployment.

Limitations and Future Scope

External factors such as news and market sentiment are not included. Future work may incorporate sentiment analysis and deep learning models like LSTM.