Cryptocurrency Tweet Sentiment Analysis: Predicting Bitcoin Volatility Using Social Media Signals

Lucas Little

University of Colorado - Boulder

CSCA 5522: Data Mining Project

Twitter Sentiment Shockwaves: Validating Historical Correlations for Bitcoin Volatility Prediction

Imagine if you could predict Bitcoin's most volatile moments by listening to what millions of Twitter users were saying.

This project explores whether social media sentiment contains statistically significant predictive signals for cryptocurrency market volatility.

We take a rigorous backtesting approach to validate historical correlations, which is particularly relevant given the influence of retail sentiment in crypto markets.

Research Question & Motivation

The Challenge:

- Cryptocurrency markets exhibit extreme volatility.
- Markets are heavily driven by retail sentiment.
- Social media creates a direct feedback loop between online sentiment and market behavior.

Research Question:

Do historical social-media patterns contain statistically significant predictors of Bitcoin volatility?

Our Approach:

- **Multi-platform sentiment analysis:** Combining Twitter data with advanced NLP models like FinBERT.
- **High-frequency analysis:** Using 15-minute temporal resolution to capture rapid sentiment shifts.
- **Historical validation:** Focusing on rigorous statistical testing rather than forward prediction.

Goal & Significance:

- Demonstrate that sentiment anomalies historically preceded significant price moves (±5% Bitcoin price movements).
- Provide a foundation for future predictive models and validate the use of social media sentiment in cryptocurrency trading strategies.

Literature Foundation

Prior Work Overview:

- **Long et al. (2025):** Studied the relationship between social media sentiment, market volatility, traders in crypto markets.
- **Youssfi Nouira et al. (2023):** Combined sentiment analysis from Twitter and Google News to predict Bitcoin price movements.
- **Brauneis and Sahiner (2024):** Compared volatility forecasting methods, sentiment-enhanced approaches, and machine learning techniques.

Research Gap:

- Lack of rigorously backtested, high-frequency, multi-platform sentiment anomalies as predictors of Bitcoin volatility.
- Most existing studies either focus on single platforms, use daily aggregations, or rely on forward-looking predictions without proper historical validation.

Data Sources

Dataset Overview:

Two primary data sources: Twitter data and Bitcoin historical price data, precisely aligned for temporal correlation analysis.

Data Alignment Challenge:

- Aligning multi-platform datasets with different temporal resolutions.
- Ensuring data quality through robust preprocessing pipelines.

Methodology

Sentiment Pipeline

- **FinBERT:** A BERT model fine-tuned on financial text to understand cryptocurrency-specific language.
- **Sentiment Aggregation:** Aggregated sentiment per 15-minute window, calculating mean, volume, momentum, and dispersion.

Anomaly Detection

- **STL Decomposition:** To identify sentiment deviations that were statistically significant relative to normal patterns.
- **Z-score Analysis:** On residuals to flag anomalies.

Temporal Correlation Analysis

Looked for ±5% Bitcoin price moves within 2 hours of identified sentiment anomalies.

Statistical Testing

Diebold-Mariano tests: To ensure results weren't due to look-ahead bias.

Evaluation Framework & Metrics

Rolling-Window Design:

Train models up to time T, then validate on the period from T to T + delta.

Prevents look-ahead bias and mimics real-world trading conditions.

Performance Metrics:

- **Precision:** Percentage of sentiment alerts that correctly predict volatility events.
- **Recall:** Percentage of actual volatility events preceded by sentiment alerts.
- **F1 Score:** Balances precision and recall (target ≥ 0.50).
- **Sharpe Ratio:** Measures risk-adjusted returns (target > 0).

Baseline Comparisons:

- 30-day realized volatility simple moving averages.
- Random alerts.
- Traditional technical indicator signals (RSI and MACD crossovers).

Timeline & Contributions

1

Week 1

Data collection and exploratory data analysis.

2

Weeks 2–3

Sentiment pipeline and feature engineering.

3

Week 4

Backtesting and baseline comparisons.

4

Week 5

Statistical testing and visualization.

5

Week 6

Final report and presentation.

Future Work

Real-Time Implementation:

If historical validation proves successful, the next step is building a live system.

Challenges include handling live data streaming, API rate limits, and data quality issues.

Potential for feedback loops where trading strategies alter sentiment patterns.

Requires robust risk management controls.