# Understanding Bitcoin Price Movements Through Behavioral Data

**DSA210 Term Project Summary**

## 1. Introduction & Motivation

This project investigates whether **behavioral and attention-based indicators** can predict **Bitcoin's short-term price movements**. Unlike traditional financial assets, Bitcoin's value is heavily influenced by **human behavior**, **social media narratives**, and **collective sentiment**, rather than intrinsic fundamentals.

As an active follower of cryptocurrency markets, I observed that online discussions, sentiment shifts, and attention spikes often shape market mood, motivating this analysis. The project combines multiple behavioral data sources — **tweet sentiment**, **Crypto Fear & Greed Index**, and **Google Trends** — to examine whether these publicly available signals contain predictive information for **next-day Bitcoin price direction**.

**Research Question:** Can publicly available behavioral indicators (social media sentiment, attention metrics, and fear–greed measures) provide predictive power for Bitcoin's next-day price direction?

## 2. Data Sources & Feature Engineering

The analysis integrates four primary datasets:

- **Bitcoin Price Data** (Yahoo Finance): daily close prices and returns
- **Bitcoin-related Tweets** (Kaggle): millions of tweets used to measure sentiment and market mood
- **Google Trends**: web and YouTube search interest for Bitcoin- and crypto-related queries
- **Crypto Fear & Greed Index**: a daily indicator of investor psychology

### Why Feature Engineering Was Necessary

Raw behavioral and price data alone cannot capture the **temporal dynamics** of financial markets. Therefore, extensive feature engineering was applied:

- **Tweet Sentiment Features:** Tweets were processed using VADER sentiment analysis to produce daily average sentiment scores, bull/bear ratios, sentiment spread, and sentiment momentum.
- **Technical Indicators:** Daily returns, 7-day volatility, momentum, and trend scores were computed, as raw price levels are not directly suitable for predictive modeling.
- **Lag Features (56 total):** Past values (1, 2, 3, and 7-day lags) were introduced to capture delayed market reactions to behavioral signals.
- **Rolling Statistics (20 total):** 7-day and 14-day rolling averages were used to smooth daily noise and identify sustained behavioral and market trends.
- **Regime & Interaction Features:** Volatility regime indicators and interaction terms (e.g., Sentiment × Volume, Trend × Volatility) were created to capture conditional effects.

## 3. Methodology

The project followed a complete data science workflow:

1. Data collection and cleaning
2. Exploratory data analysis (EDA) with visualizations
3. Statistical hypothesis testing: Pearson correlation tests and Mean-difference tests (t-tests and Mann–Whitney U tests)
4. Machine learning classification using Logistic Regression and Random Forest

The objective was to test whether behavioral indicators provide predictive power for **next-day Bitcoin price direction**.

# 4. Key Results

### 4.1 Hypothesis Testing Results

**Correlation Tests (Pearson)**

| Test | r value | p-value | Result |
|---|---|---|---|
| Sentiment vs Next Return | -0.053 | 0.657 | Fail to reject $H_0$ |
| Sentiment Spread vs Next Return | 0.070 | 0.555 | Fail to reject $H_0$ |
| Fear & Greed vs Next Return | 0.033 | 0.780 | Fail to reject $H_0$ |

All p-values exceed 0.05, indicating **no statistically significant linear relationship** between behavioral indicators and next-day Bitcoin returns.

**Mean Difference Tests (Up Days vs Down Days)**

Independent sample t-tests and Mann–Whitney U tests were used to compare behavioral variables on up versus down days. Most features showed **no significant differences** ($p > 0.05$). Only **Sentiment Spread** produced a marginally significant t-test result ($p = 0.031$), which was **not confirmed** by the non-parametric Mann–Whitney test ($p = 0.066$), suggesting a weak and unstable signal.

### 4.2 Machine Learning Results

| Metric | Logistic Regression | Random Forest |
|---|---|---|
| Accuracy | 48.92% | 52.52% |
| Precision | 45.71% | 48.78% |
| Recall | 49.23% | 30.77% |
| F1-Score | 47.41% | 37.74% |

For reference, a naive random baseline would achieve approximately **50% accuracy**. Both models perform **at or near this baseline**, indicating that they do not extract meaningful predictive signals beyond random chance. Feature importance scores were **low and evenly distributed**, with top contributors such as fng_value_lag3 (≈3.3%) and Volatility_7d_lag7 (≈2.5%), further suggesting the absence of dominant predictive patterns.

# 5. Conclusion

Do behavioral indicators predict Bitcoin's short-term price direction? **The evidence strongly suggests no.**

Both statistical hypothesis tests and machine learning models failed to identify robust or exploitable relationships between behavioral indicators and next-day Bitcoin price movements. Importantly, this project demonstrates that **the absence of predictive power is itself an informative result**, emphasizing the difficulty of exploiting publicly available sentiment data in highly liquid and information-efficient markets.

These findings support the **Efficient Market Hypothesis** in the context of cryptocurrencies: behavioral signals that are widely observable appear to be rapidly incorporated into prices. Short-term Bitcoin returns are dominated by **noise, rapid information diffusion, and market microstructure effects**, limiting the usefulness of daily sentiment-based predictors.

## 6. Limitations & Future Work

**Limitations:**

• Tweet data may contain noise, bots, and repeated content
• Google Trends required interpolation due to missing dates
• Only next-day prediction horizons were tested
• External factors such as macroeconomic news and on-chain activity were not modeled

**Future Work:**

• Extend prediction horizons to 3-day or 7-day returns
• Incorporate on-chain metrics (e.g., transaction volume, whale activity)
• Apply deep learning approaches (e.g., LSTM models)
• Conduct event-based analyses around major market news

## 7. AI Assistance Disclosure

In accordance with academic integrity requirements, AI assistance was used in two limited and well-defined areas:

**1. Chunk-Based Tweet Processing:** Due to the large size of the raw tweet dataset (~2GB+), AI assistance was used to design and implement an efficient chunk-based processing pipeline for sentiment extraction.

**2. Machine Learning Pipeline Development:** AI guidance was used to structure the machine learning workflow, including feature engineering strategies (lag features, rolling statistics, regime indicators) and model evaluation design.

**Examples of AI Prompts Used**

“Design a memory efficient Python pipeline to process a multi-gigabyte tweet dataset using chunk-based sentiment analysis.”

“Suggest a clean and modular structure for a machine learning pipeline with lag features, rolling statistics, and time-aligned targets.”

The AI-generated outputs consisted of:

- High-level implementation suggestions and code structure ideas

- Optimization strategies for large-scale data processing

- General guidance on feature engineering and model evaluation