

Highlights

Análisis del Impacto del Sentimiento en Twitter sobre el Precio de Bitcoin (2021–2023)

Katherine López Ramírez

- Análisis de 4,69 millones de tuits relacionados con Bitcoin (2021–2023) usando filtros de relevancia de cuenta.
- Clasificación de sentimiento con RoBERTa adaptado a redes sociales y evaluación multiescala (día, hora y minuto).
- Twitter funciona como *termómetro social* del mercado: refleja dinámicas colectivas más que predecir precios de forma directa.
- En subidas extremas, un sentimiento positivo acumulado en 3–7 días previos se asocia con mayores retornos.
- En caídas extremas, predominó un optimismo engañoso antes de desplomes, mostrando el valor del sentimiento como señal contextual.

Análisis del Impacto del Sentimiento en Twitter sobre el Precio de Bitcoin (2021–2023)

Katherine López Ramírez^a

^a*Universidad Europea de Madrid, Madrid, España*

Abstract

Este estudio analiza la relación entre el sentimiento expresado en Twitter y el precio de Bitcoin durante 2021–2023, a partir de 4,69 millones de tuits. Se aplicaron filtros de relevancia de cuenta (verificadas, con más de 10.000 y con más de 1.000.000 de seguidores) y se clasificó el sentimiento con el modelo RoBERTa adaptado a redes sociales.

El análisis multiescala (día, hora, minuto) muestra que, en general, la correlación entre sentimiento y retornos de Bitcoin es baja, lo que confirma que Twitter funciona más como *termómetro social* que como predictor aislado. Sin embargo, en episodios extremos se identificaron patrones relevantes: en las subidas más fuertes, el sentimiento positivo acumulado en los días previos estuvo asociado a mayores retornos, mientras que en las caídas extremas predominaba un optimismo engañoso que antecedió a desplomes inesperados.

Estos hallazgos aportan evidencia empírica a un debate abierto en la literatura, mostrando que el sentimiento en redes sociales refleja dinámicas colectivas y narrativas del mercado, con capacidad limitada de predicción general, pero con señales útiles en contextos de alta volatilidad.

1. Introduction

El mercado de las criptomonedas ha experimentado un crecimiento acelerado en la última década, consolidándose como un activo financiero de interés tanto para inversores individuales como para instituciones. Entre ellas, Bitcoin se ha posicionado como la criptomoneda más representativa, caracterizada por su elevada volatilidad y su sensibilidad a factores externos como regulaciones, avances tecnológicos o noticias globales. Esta alta volatilidad ha motivado la búsqueda de modelos capaces de explicar o anticipar sus fluctuaciones.

En paralelo, las redes sociales se han convertido en un canal clave para la difusión de noticias, opiniones y predicciones sobre mercados financieros. En particular, Twitter ha emergido como un espacio donde inversores, analistas y medios especializados comparten en tiempo real información que puede influir en las expectativas del mercado. De este modo, el análisis de sentimientos (*sentiment analysis*) aplicado a Twitter se presenta como una herramienta relevante para explorar la relación entre las emociones colectivas y los movimientos de precios en el mercado de criptomonedas.

Estudios previos han mostrado resultados diversos. Lupu [1] encontró que el sentimiento en Twitter afecta retornos, liquidez y volatilidad de Bitcoin en el corto plazo. Ulu y Ulu [2] concluyeron que dicha relación tiende a ser débil y no sostenida en horizontes largos. Pathmanathan [3] evidenció que el sentimiento negativo puede anticipar caídas cuando participan figuras influyentes como Elon Musk, mientras que Cary [5] observó un comportamiento colectivo (*herding*) de tono neutral tras el *crash* de mayo de 2022. Desde un enfoque metodológico, Zou y Herremans [6] desarrollaron el modelo *PreBit* que combina variables lingüísticas y técnicas para predecir movimientos extremos, y D’Amario y Ciganovic [7] utilizaron modelos VAR interpretables para identificar relaciones bidireccionales entre sentimiento y precio. Asimismo, Akbiyik et al. [8] destacaron que la volatilidad puede predecirse mejor considerando quién tuitea que el contenido de los mensajes.

De forma complementaria, investigaciones recientes en *Blockchain: Research and Applications* han ampliado el debate hacia factores técnicos, sociales y de percepción digital. Bouraga [12] mostró que en Ethereum la popularidad depende tanto de variables de red como de la actividad en Twitter. Bakhtiar et al. [13] concluyeron que los sentimientos de mercado (p. ej., miedo y codicia, búsquedas en Google) pesan más que los fundamentos técnicos en la valoración de criptomonedas. Chalkiadakis et al. [14] demostraron que las noticias negativas ejercen efectos más persistentes sobre los precios que las positivas. Ben Hamadou et al. [15] evidenciaron que el sentimiento medido con Google Trends afecta a Bitcoin, Ethereum y Litecoin en periodos de crisis, mientras que Tether se mantiene estable. Finalmente, Ghazouani et al. [16] encontraron que el sentimiento negativo en Reddit influye fuertemente en Bitcoin, mientras que Ethereum responde más a factores tecnológicos.

En este contexto, nuestro estudio contribuye al debate evaluando el impacto del sentimiento en Twitter sobre el precio de Bitcoin durante el periodo 2021–2023. La principal novedad radica en el filtrado de relevancia de las cuentas (≥ 10.000 y $\geq 1.000.000$ seguidores) y en la evaluación sistemática de

correlaciones en distintas escalas temporales (día, hora y minuto). De este modo, se busca determinar en qué medida el sentimiento expresado en redes sociales se asocia con la dinámica de los retornos de Bitcoin.

2. Methodology

La metodología se diseñó para evaluar de manera rigurosa la relación entre el sentimiento en Twitter y la evolución del precio de Bitcoin durante el periodo 2021–2023. El proceso combinó la construcción de un dataset masivo de tuits, su clasificación con técnicas de aprendizaje profundo y el análisis cuantitativo de su asociación con los retornos de Bitcoin en diferentes escalas temporales.

2.1. Obtención y filtrado de datos

Se recopiló un total de 4.689.354 tuits relacionados con Bitcoin mediante búsquedas basadas en palabras clave y *hashtags* asociados. Para garantizar la relevancia de la muestra y reducir el ruido informativo, se establecieron dos escenarios de filtrado:

- **Escenario 1 (base amplia):** tuits de cuentas verificadas o con al menos 10.000 seguidores ($n = 398.877$).
- **Escenario 2 (influencia elevada):** tuits de cuentas verificadas con un millón o más de seguidores ($n = 5.952$).

En ambos casos se eliminaron los retuits para evitar duplicidad y sobreponderación de mensajes replicados.

2.2. Preprocesamiento del texto

El contenido de los tuits fue normalizado mediante: conversión a minúsculas, eliminación de enlaces, menciones y caracteres especiales, y conservación de *hashtags* asociados a Bitcoin. El resultado fue un campo limpio (`clean_text`) empleado como insumo para el análisis de sentimientos.

2.3. Clasificación de sentimiento

El sentimiento de cada tuit se clasificó con el modelo *Twitter RoBERTa-base* de CardiffNLP [9, 10], optimizado para lenguaje en redes sociales. La implementación se realizó con la librería **Transformers** [11]. Cada mensaje se etiquetó como positivo, negativo o neutral, registrando además la probabilidad asociada (`sentiment_score`), lo que permitió la construcción de métricas agregadas más robustas.

2.4. Variables de sentimiento

A partir de las clasificaciones individuales se generaron métricas en distintos horizontes temporales:

- Distribución de tuits por polaridad (positivo, negativo, neutral).
- Sentimiento neto:

$$\text{Sent_Net} = \text{Positivos} - \text{Negativos}$$

- Sentimiento neto normalizado por volumen total de tuits.
- Promedios móviles de 3 y 7 días del sentimiento firmado (`score_signed`).
- Ratio de sentimiento: proporción de tuits positivos frente a negativos.

2.5. Datos financieros

Los precios históricos de Bitcoin se obtuvieron de fuentes confiables (Blockchain.com, Binance y Yahoo Finance), empleando precios de cierre ajustados. A partir de ellos se calcularon los retornos logarítmicos diarios:

$$r_t = \ln \left(\frac{P_t}{P_{t-1}} \right).$$

2.6. Análisis de correlación y eventos extremos

La relación entre sentimiento y retornos se evaluó en tres escalas: diaria, horaria y minutal. Se calcularon correlaciones de Pearson, Spearman y Kendall en diferentes escenarios:

- Correlación contemporánea (mismo día/hora/minuto).
- Correlación con rezagos positivos (sentimiento adelantado) y negativos (sentimiento rezagado).
- Ventanas móviles de hasta ± 7 días y ± 60 minutos.

Adicionalmente, se identificaron los 20 eventos extremos más relevantes (10 subidas y 10 caídas de mayor magnitud en 2021–2023). Para cada evento se analizaron las métricas de sentimiento acumuladas en los días previos, con el fin de explorar patrones específicos de euforia o pánico colectivo.

2.7. Visualización

Se elaboraron gráficos para representar:

- Evolución del volumen de tuits por polaridad.
- Comparación entre sentimiento neto y precio de Bitcoin.
- Correlaciones por rezago entre métricas de sentimiento y retornos.
- Picos de sentimiento en relación con eventos extremos de mercado.

Estas visualizaciones facilitaron la interpretación de patrones temporales y el contraste con los hallazgos cuantitativos.

3. Results

En esta sección se presentan los principales hallazgos del análisis de 4,69 millones de tuits y su relación con el precio de Bitcoin entre 2021 y 2023. Los resultados combinan una visión descriptiva de la dinámica del sentimiento en Twitter con un análisis cuantitativo de correlaciones y eventos extremos.

3.1. Distribución del sentimiento

El análisis inicial muestra que la mayoría de los tuits corresponde a la categoría neutral (65–70%), mientras que los positivos representan en promedio un 20–25% y los negativos un 10–12%. Este predominio del tono neutral confirma que, aunque Twitter refleja episodios de polarización, el discurso general sobre Bitcoin tiende a ser informativo o descriptivo más que emocional (Figura 1).

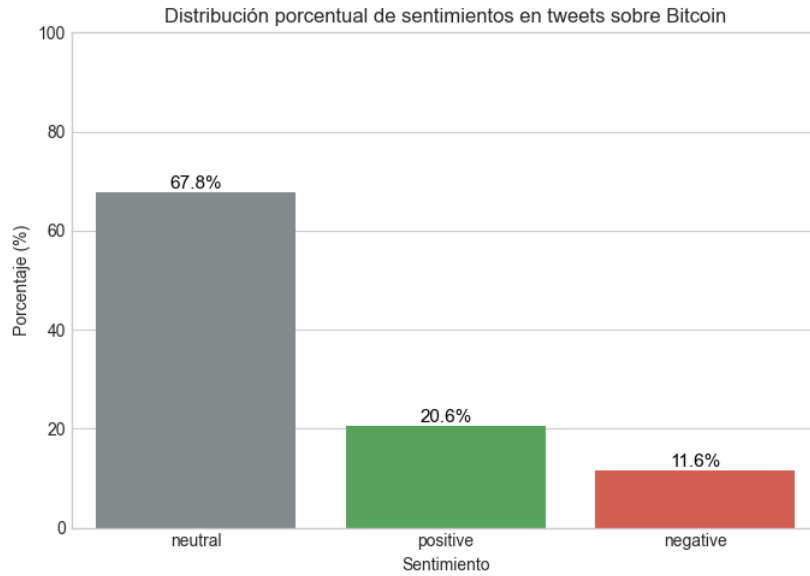


Figure 1: Distribución de tuits clasificados en positivos, negativos y neutrales.

3.2. Volumen y dinámica temporal de la actividad en Twitter

El volumen diario de tuits exhibe picos claramente asociados a episodios de alta volatilidad del mercado, como la compra de Bitcoin por parte de Tesla en febrero de 2021, la caída del mercado en mayo de 2022 o el colapso de FTX en noviembre de 2022. Estos resultados confirman que Twitter funciona como un *termómetro social* del ecosistema cripto, reflejando en tiempo real la intensidad de la conversación colectiva frente a eventos de impacto (Figura 2).

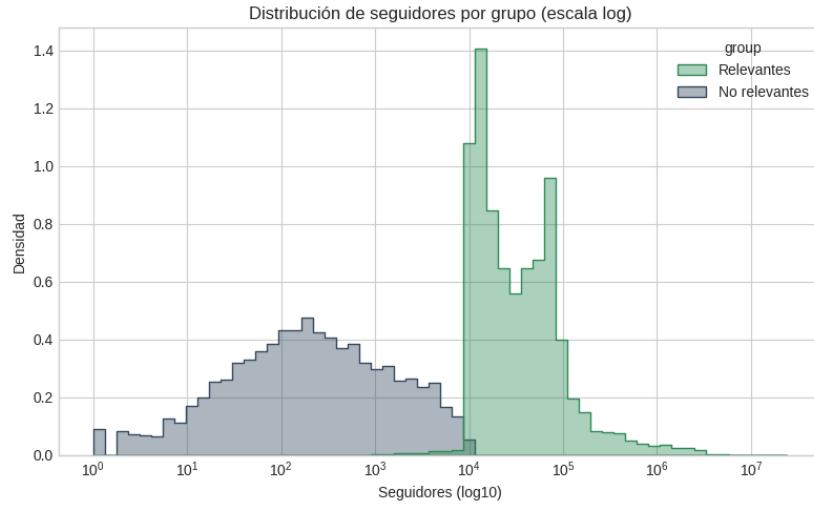


Figure 2: Evolución temporal del volumen de tuits relacionados con Bitcoin.

3.3. Evolución comparativa entre sentimiento y precio de Bitcoin

La comparación entre el sentimiento neto y la serie de precios de Bitcoin revela paralelismos puntuales en episodios de estrés de mercado, aunque sin un patrón consistente para todo el periodo de estudio (Figura 3). Esto sugiere que el sentimiento no constituye un predictor directo de los retornos diarios, pero sí captura las narrativas dominantes que acompañan los grandes movimientos de mercado.

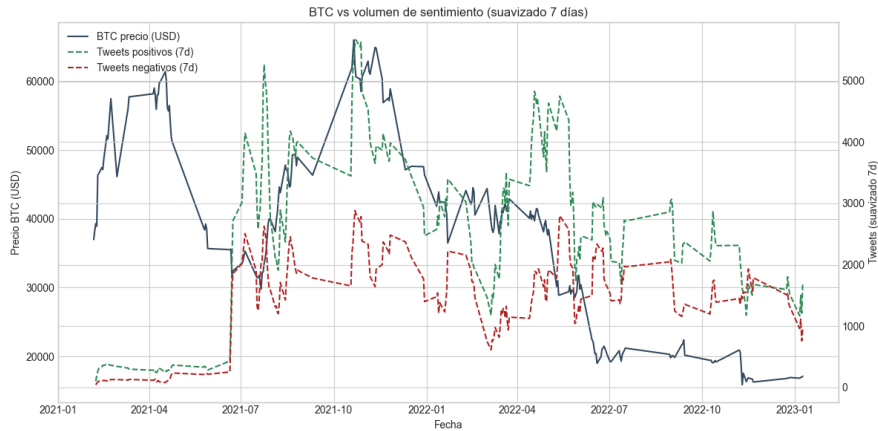


Figure 3: Comparación entre el sentimiento neto agregado y el precio de Bitcoin.

3.4. Correlaciones en diferentes escalas temporales

Los coeficientes de Pearson calculados en horizontes diario, horario y minutal muestran correlaciones bajas en general (Tabla 1), lo que confirma que el sentimiento en Twitter no anticipa de manera sistemática los retornos. Sin embargo, a nivel horario se identificó una correlación moderada (0.105) con un rezago de -4 horas, indicando que ciertos impulsos de sentimiento pueden asociarse con movimientos intradía.

Table 1: Correlación entre sentimiento neto y retornos de Bitcoin

Horizonte temporal	Máxima correlación observada	Rezago identificado
Diario	0.02	Ninguno significativo
Horario	0.105	-4 horas
Minutal	0.016	No consistente

3.5. Patrones en eventos extremos

Un aporte central del estudio fue el análisis de los 20 eventos más extremos (10 subidas y 10 caídas). Los resultados muestran un patrón temporal diferenciado entre escenarios alcistas y bajistas:

- En las **subidas extremas**, el sentimiento positivo acumulado en ventanas de 3–7 días previos presentó correlaciones moderadas con la magnitud del alza (ej. `sent_ratio_3d` ≈ 0.56). Este resultado indica que un optimismo sostenido durante varios días puede anticipar rallies significativos, como ocurrió el 8 de febrero de 2021 tras el anuncio de la compra de 1,5 mil millones de USD en Bitcoin por parte de Tesla. En este caso, la narrativa positiva en redes estuvo alineada con la noticia institucional, potenciando el impulso alcista.
- En las **caídas extremas**, en cambio, el sentimiento previo también era predominantemente positivo, pero concentrado en ventanas más cortas de 1–3 días. Este optimismo engañoso antecedió desplomes inesperados, como el 21 de enero de 2022, cuando 128 tuits netos positivos coincidieron con un ambiente de confianza que se vio abruptamente revertido por el endurecimiento de la política monetaria de la Fed. De forma similar, en otras caídas fuertes de 2022 el sentimiento no reflejaba riesgos inminentes, lo que resalta la desconexión entre la percepción social y los shocks externos.

Un caso particular lo constituye el colapso de FTX en noviembre de 2022: a diferencia de otras caídas, en este evento el sentimiento ya era netamente negativo en los días previos (−60 tuits netos), reflejando que la crisis era percibida con anticipación en la conversación digital.

En síntesis, el sentimiento positivo acumulado en redes sociales refleja un estado de alta expectativa colectiva. En periodos de 3–7 días, este optimismo se asocia a subidas cuando está respaldado por noticias favorables; en horizontes de 1–3 días, en cambio, puede preceder caídas inesperadas cuando las narrativas en redes no logran captar señales de riesgo provenientes del entorno macroeconómico o regulatorio.

4. Discussion

Los hallazgos de este estudio confirman que el sentimiento expresado en Twitter no constituye un predictor sistemático de los retornos de Bitcoin en horizontes diarios, horarios o minutales. Sin embargo, aportan nueva evidencia de su papel como *termómetro social* del mercado: la conversación digital refleja la intensidad emocional de la comunidad y acompaña, aunque no siempre anticipe, los grandes movimientos de precio.

En términos agregados, las correlaciones entre sentimiento y retornos fueron bajas, en línea con trabajos como Lupu [1] y Ulu y Ulu [2], quienes concluyen que la influencia del sentimiento sobre el precio de Bitcoin tiende a ser débil y transitoria. De forma similar, Pathmanathan [3] y Cary [5] mostraron que la dinámica del sentimiento depende en gran medida de episodios puntuales y de la intervención de figuras influyentes, más que de una relación estadística estable en el tiempo.

La principal contribución de nuestro estudio radica en el análisis de **eventos extremos**. A diferencia de estudios previos, identificamos que la acumulación de sentimiento positivo presenta un comportamiento diferenciado según el contexto:

- En subidas extremas, el optimismo acumulado durante 3–7 días previos estuvo asociado a retornos significativamente más altos, especialmente cuando la narrativa digital estuvo respaldada por noticias favorables, como el caso de la compra de Bitcoin por parte de Tesla en febrero de 2021.
- En caídas extremas, por el contrario, predominó un optimismo engañoso en ventanas cortas de 1–3 días, lo que refleja un exceso de

confianza previo a desplomes inesperados. Ejemplo de ello fue la caída del 21 de enero de 2022, cuando abundaban tuits positivos pese al endurecimiento de la política monetaria de la Fed. Solo en casos de crisis ampliamente percibidas, como el colapso de FTX en noviembre de 2022, el sentimiento negativo se adelantó a la caída.

Este hallazgo amplía la literatura de *Blockchain: Research and Applications*. Bouraga [12] mostró que la popularidad de Ethereum depende tanto de variables de red como de la actividad en Twitter; Bakhtiar et al. [13] señalaron que los sentimientos colectivos pesan más que los fundamentos técnicos en la valoración de criptomonedas; Chalkiadakis et al. [14] evidenciaron que las noticias negativas tienen efectos más persistentes que las positivas; Ben Hamadou et al. [15] demostraron la importancia del sentimiento en periodos de crisis usando Google Trends; y Ghazouani et al. [16] destacaron que Bitcoin es particularmente sensible al pesimismo social frente a Ethereum. Nuestro estudio complementa estas perspectivas al mostrar que el sentimiento positivo, aunque no predice consistentemente el mercado, sí permite caracterizar estados de euforia o confianza que, combinados con noticias exógenas, pueden anticipar tanto rallies como desplomes.

En síntesis, Twitter no debe interpretarse como un predictor autónomo del precio de Bitcoin, sino como un indicador contextual de la intensidad emocional del mercado. El valor del sentimiento radica en su capacidad para capturar narrativas colectivas que, al interactuar con factores externos (regulatorios, institucionales o macroeconómicos), condicionan la dirección y magnitud de los movimientos extremos.

Table 2: Comparación de estudios previos y el presente trabajo

Estudio / Año	Dataset	Metodología / Tecnología	Conclusiones principales
Bouraga (2023) [12]	Tweets sobre Ethereum, métricas de red	Análisis de popularidad combinando factores de red y actividad en Twitter	La popularidad de Ethereum depende tanto de variables técnicas como de la interacción social en redes.
Bakhtiar et al. (2023) [13]	Datos de mercado + sentimiento (miedo/codicia, búsquedas en Google)	Modelos econométricos y análisis de sentimiento macro	Los sentimientos colectivos pesan más que los fundamentos técnicos en la valoración de criptomonedas.
Chalkiadakis et al. (2022) [14]	Noticias y datos de mercado de criptomonedas	Modelos de causalidad estadística (VAR)	Las noticias negativas tienen efectos más persistentes sobre precios que las positivas.
Ben Hamadou et al. (2024) [15]	Google Trends (Bitcoin, Ethereum, Litecoin, Tether)	Modelos de causalidad en cuantiles	El sentimiento digital afecta retornos en periodos de crisis; Tether se mantiene estable.
Ghazouani et al. (2025) [16]	Publicaciones en Reddit	Análisis de sentimiento + modelos econométricos	Bitcoin es más sensible al pesimismo social, mientras que Ethereum responde más a factores tecnológicos.
Presente estudio (2025)	4.69M tuits sobre Bitcoin (2021–2023). Escenarios: 10k y 1M seguidores.	RoBERTa para sentimiento, métricas agregadas (neto, ratio, score firmado), correlaciones multiescala, análisis de 20 eventos extremos	Correlaciones bajas en general, pero hallazgos clave: (i) optimismo acumulado en 3–7 días se asocia a subidas con noticias favorables (Tesla, 2021); (ii) optimismo en 1–3 días previos puede anteceder caídas inesperadas (Fed, 2022); (iii) en crisis percibidas (FTX, 2022), el sentimiento negativo anticipa caídas. Twitter funciona como <i>termómetro social</i> .

5. Conclusions and Future Work

Este estudio examinó la relación entre el sentimiento expresado en Twitter y el precio de Bitcoin durante el periodo 2021–2023, aplicando un procesamiento riguroso de más de 4,6 millones de tuits y filtrando por relevancia de cuentas. Los resultados muestran que, aunque existe cierta asociación entre sentimiento y retornos, la correlación es baja y no estadísticamente significativa en la mayoría de horizontes temporales (diario, horario y por minuto).

Estos hallazgos permiten concluir que el sentimiento en Twitter no constituye un predictor fiable del precio de Bitcoin, sino que refleja de manera más bien descriptiva la percepción colectiva de la comunidad frente a los movimientos del mercado. En otras palabras, funciona más como un indicador de acompañamiento que como una señal de predicción.

Entre las principales aportaciones de este trabajo destacan:

- Un proceso sistemático de filtrado de tuits basado en la relevancia de la cuenta ($\geq 10k$ y $\geq 1M$ seguidores).
- La aplicación de un modelo de *sentiment analysis* entrenado específicamente en redes sociales (RoBERTa).
- El análisis multiescala (día, hora, minuto) de la relación entre sentimiento neto y retornos de Bitcoin.

Como líneas futuras de investigación, se propone:

- Adoptar un enfoque inverso: identificar episodios de alta volatilidad (grandes caídas o subidas) y analizar qué patrones de sentimiento los precedieron.
- Ampliar la base de datos con fuentes adicionales de redes sociales (p. ej., Reddit, foros especializados) para obtener una visión más integral.
- Evaluar métodos más sofisticados de modelado, como pruebas de causalidad (Granger) o modelos VAR, que permitan explorar relaciones más allá de la simple correlación.
- Incorporar métricas de volatilidad y liquidez para estudiar el papel del sentimiento en la dinámica más amplia del mercado de criptomonedas.

References

- [1] Lupu, I. (2025). *Cryptocurrency Efficiency and Twitter Sentiment: An Analysis of the Impact on Return, Liquidity, and Volatility*. MDPI Data.
- [2] Ulu, Ç., & Ulu, C. (2024). *Examining the Relationship Between Twitter Sentiment and Bitcoin Prices*. Beta Journal.
- [3] Pathmanathan, M., et al. (2022). *Bitcoin Price Fluctuation by Twitter Sentiment Analysis Pre- and Post-COVID-19 Pandemic*. ResearchGate.
- [4] Jahanbin, M., et al. (2024). *Cryptocurrency Tweets Sentiment Dataset (2021–2023)*. BMC Research Notes.
- [5] Cary, R. (2024). *Twitter Herding After the May 2022 Bitcoin Crash*. Journal of Financial Innovation.
- [6] Zou, L., & Herremans, D. (2022). *PreBit: A Multimodal Model for Bitcoin Price Prediction*. arXiv:2206.00648.
- [7] D’Amario, F., & Ciganovic, I. (2022). *Causal Impact of Social Media Sentiment on Bitcoin Price*. arXiv:2210.00883.
- [8] Akbiyik, A., et al. (2021). *Predicting Volatility in Bitcoin Using Twitter Metadata*. arXiv:2110.14317.
- [9] Barbieri, F., Camacho-Collados, J., Espinosa-Anke, L., & Neves, L. (2020). *TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification*. Findings of EMNLP 2020.
- [10] Camacho-Collados, J., Scarlini, B., Barbieri, F., Neves, L., et al. (2022). *TweetNLP: Cutting-Edge Natural Language Processing for Social Media*. arXiv:2206.14774.
- [11] Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., et al. (2020). *Transformers: State-of-the-Art Natural Language Processing*. Proceedings of EMNLP 2020: System Demonstrations.
- [12] Bouraga, S. (2023). Assessing the impact of network factors and Twitter data on Ethereum’s popularity. *Blockchain: Research and Applications*, Elsevier.

- [13] Bakhtiar, M., Luo, J., & Adelopo, I. (2023). The impact of fundamental factors and sentiments on the valuation of cryptocurrencies. *Blockchain: Research and Applications*, Elsevier.
- [14] Chalkiadakis, I., Zaremba, A., Peters, G., & Chantler, M. (2022). On-chain analytics for sentiment-driven statistical causality in cryptocurrencies. *Blockchain: Research and Applications*, Elsevier.
- [15] Ben Hamadou, S., Mezghani, I., & Boujelbène Abbes, M. (2024). Time-varying nexus and causality in the quantile between Google investor sentiment and cryptocurrency returns. *Blockchain: Research and Applications*, Elsevier.
- [16] Ghazouani, I., Ghazouani, I., & Omri, A. (2025). Virtual influence, real impact: Deciphering social media sentiment and its effects on cryptocurrency market dynamics. *Blockchain: Research and Applications*, Elsevier.