Market-Aware Event Timeline Summarization: Integrating Price Signals to Improve Financial News Understanding

Heesoo Park, Sujin Park, Dongjun Lee
Dunamu Inc.
{belle, avy, tonny}@dunamu.com

Abstract

We present a practical system that supports in-depth analysis of cryptocurrency markets through timeline-based event detection and contextual summarization. Our framework processes continuous news streams, identifies price-relevant events, and organizes them into semantic timelines with concise background summaries generated by large language models (LLMs). This design allows traders and analysts to retrospectively explore events alongside price charts, facilitating a deeper understanding of how news developments relate to market fluctuations. By transforming unstructured news data into structured insights, the system provides a valuable tool for market analysis, risk evaluation, and behavioral studies in volatile trading environments. A live demonstration of our system is available at: https://uptrace.duml.io

Introduction

Cryptocurrency markets are highly volatile, with prices frequently influenced by breaking news, regulatory decisions, economic trends, and influential statements from key figures (Sun, Liu, and Sima 2020; Alizadeh et al. 2024). Traders and analysts, however, face difficulties in determining whether a new headline represents an isolated occurrence or a continuation of an ongoing trend (Garcia and Schweitzer 2015). This lack of structured context hinders systematic analysis and complicates the assessment of event significance (Kim, Goetzmann, and Shiller 2023).

To address this challenge, we introduce a timeline-based event extraction and summarization system that links market volatility with structured event insights. The system identifies cryptocurrency-related events from heterogeneous news streams, generates concise phrases and summaries, and clusters semantically related items into coherent timelines. Daily market data is converted into return- and volatility-based textual descriptions to establish causal links between price changes and events. Events are scored for relevance, clustered, and filtered by volatility impact using a majority-voting strategy (Trad and Chehab 2025), with at most three key events per day retained. Background summaries are constructed from full news bodies, enabling users to understand not only what happened but also why it matters in historical context.

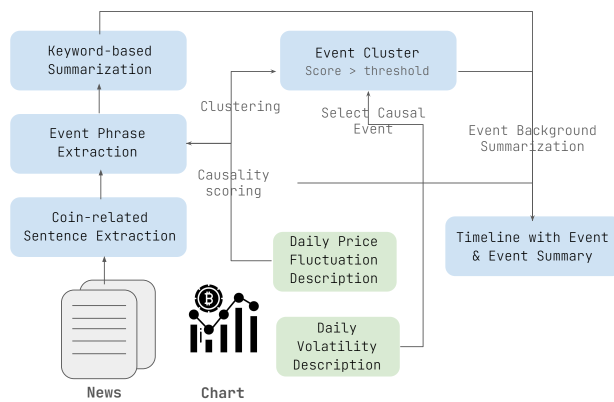

Figure 1: Overview of our event detection and summarization framework.

Unlike prior financial news visualization systems (Saxena, Janzen, and Maass 2024; Qorib, Hu, and Ng 2024; Faghihi et al. 2022), our approach directly integrates event timelines with candlestick charts, ensuring that market participants can trace asset movements back to concrete news events. This design provides richer situational awareness by combining numerical trends with contextual narratives, offering both high-level overviews and fine-grained insights. The user interface complements this workflow with an interactive design. This design minimizes cognitive load and helps users focus on the developments most relevant to their analytical objectives.

System Architecture

User Interface Figure 2 shows an intuitive interface of our system that connects price volatility with relevant events through three main components. Our tool enables users to quickly identify what happened on days with high price volatility (Parkinson 1980) for each coin by displaying markers directly on the price chart. When the user hovers over a marker, a compressed summary of the corresponding event is shown. If the user wants more details, clicking the marker redirects them to the timeline at the corresponding

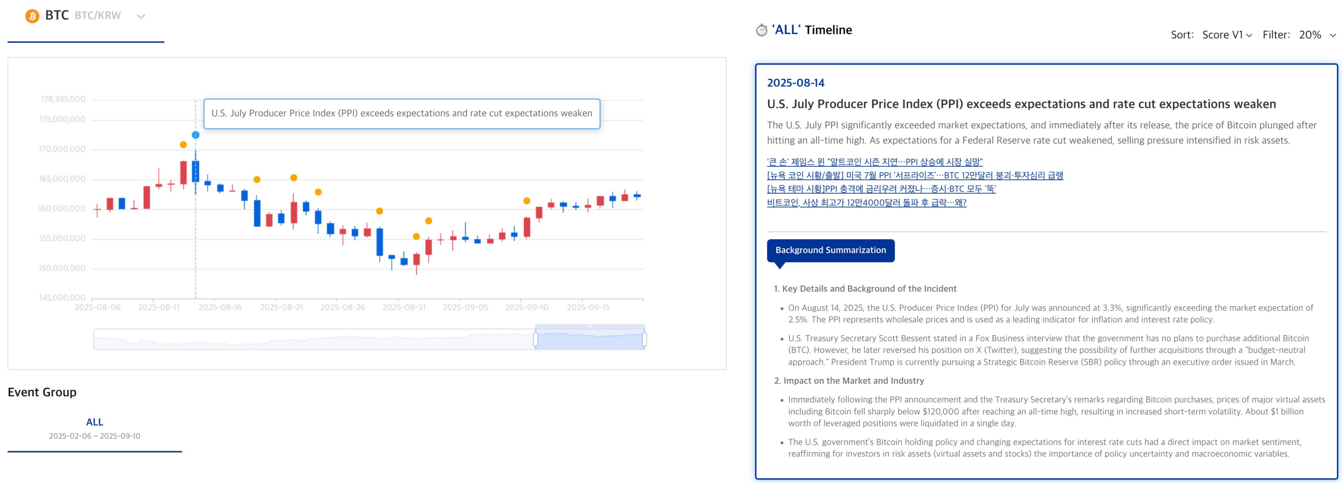

Figure 2: A screenshot of the LLM-powered event detection and timeline-based summarization system. The interface displays a BTC daily candlestick chart, where (1) milestone events are directly overlaid on the daily candlestick chart as interactive markers. Hovering over a marker reveals key event phrases, allowing users to quickly grasp the nature of market-moving news at specific time points. (2) Next to the chart, a scrollable event list presents detected events that can be sorted in reverse chronological order or by volatility impact. (3) For each event, the system provides a curated list of related news clusters.

event. Each event in the timeline provides an event phrase, an event summary, and background knowledge necessary to understand the event (Duan et al. 2024; Zhang, Xu, and Qi 2023). To help users focus on the most important events, the system allows filtering to display only the top $N\%$ of events based on volatility. The timeline can also be sorted either by volatility level or by chronological order.

Workflow Figure 1 illustrates the overall workflow of the system. When news articles are ingested, the system first performs automatic extraction of coin-specific mentions. Based on the extracted content, the system generates an event phrase and produces a corresponding event summary.

In parallel, daily market data is processed to generate textual descriptions from two perspectives: *volatility-based* and *return-based*. The volatility description is derived from the daily high and low prices, while the return-based description is calculated from the open and close prices. Causal links between these textual market descriptions and detected events are then established.

For each day, return-based movements are initially associated with news articles by assigning a relevance score to each news item. Articles that exceed a predefined threshold are retained, and their event phrases and summaries are clustered to form coherent event groups. Relevance scoring is conducted through a chain-of-thought (CoT) (Wei et al. 2023) prompting strategy. To ensure stability and reduce variance, the scoring process is repeated five times per article, and the final relevance score is computed as the average of these outputs. Our prompt design was informed by methodologies presented in the following studies. (Pratapa, Small, and Dreyer 2023; Romanou et al. 2023; Hu, Moon, and Ng 2024)

Next, the volatility-based analysis is used to identify the events with the strongest impact. A majority-voting strategy (Trad and Chehab 2025) is applied to select the most influential events. For these selected events, the system collects the full body of all related news articles and generates a detailed background summary to provide richer context.

If the selected event does not contain the top-scoring news article from the initial return-based scoring step, that article is added to the event cluster to preserve consistency. Finally, the system retains at most three key events per day to ensure conciseness and interpretability in the timeline.

Implementation All language components of our system are implemented using OpenAI's `GPT-4o-08-06` (OpenAI et al. 2024) as the backbone large language model with the temperature parameter set to 0 to ensure deterministic and consistent outputs across inference runs. For clustering, we use OpenAI's `text-embedding-3-large` model to encode news title and employ hierarchical clustering (Cohen-Addad et al. 2017). Our news pool is constructed from Upbit [1] contract news, covering major developments relevant to the cryptocurrency market. All cryptocurrency price data was retrieved via the Upbit API. [2]

Social Impact

The system enhances transparency in cryptocurrency markets by linking volatility with contextualized event summaries, enabling traders, researchers, and regulators to better understand the drivers of sudden price movements while reducing information asymmetry.

[1]https://upbit.com/
[2]https://docs.upbit.com/kr

References

Alizadeh, M.; Asgari, Y.; Samei, Z.; Yari, S.; Dehghani, S.; Kubli, M.; Zare, D.; Bermeo, J. D.; Batzdorfer, V.; and Gilardi, F. 2024. Exploring Relationships Between Cryptocurrency News Outlets and Influencers' Twitter Activity and Market Prices. In *International Conference on Advances in Social Networks Analysis and Mining*, 455–471. Springer.

Cohen-Addad, V.; Kanade, V.; Mallmann-Trenn, F.; and Mathieu, C. 2017. Hierarchical Clustering: Objective Functions and Algorithms. arXiv:1704.02147.

Duan, J.; Kou, F.; Wang, Z.; and Wei, Y. 2024. When echoes surpass voices: Market reaction to forwarded news. *International Review of Financial Analysis*, 96: 103579.

Faghihi, H. R.; Alhafni, B.; Zhang, K.; Ran, S.; Tetreault, J.; and Jaimes, A. 2022. CrisisLTLSum: A Benchmark for Local Crisis Event Timeline Extraction and Summarization. arXiv:2210.14190.

Garcia, D.; and Schweitzer, F. 2015. Social signals and algorithmic trading of Bitcoin. *Royal Society open science*, 2(9): 150288.

Hu, Q.; Moon, G.; and Ng, H. T. 2024. From Moments to Milestones: Incremental Timeline Summarization Leveraging Large Language Models. In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 7232–7246. Bangkok, Thailand: Association for Computational Linguistics.

Kim, D.; Goetzmann, W.; and Shiller, R. 2023. Economic Narratives Shape How Investors Perceive Risks. Technical report, Office of Financial Research, US Department of the Treasury.

OpenAI; :; Hurst, A.; Lerer, A.; Goucher, A. P.; Perelman, A.; Ramesh, A.; Clark, A.; Ostrow, A.; Welihinda, A.; Hayes, A.; Radford, A.; Madry, A.; Baker-Whitcomb, A.; Beutel, A.; Borzunov, A.; Carney, A.; Chow, A.; Kirillov, A.; Nichol, A.; Paino, A.; Renzin, A.; Passos, A. T.; Kirillov, A.; Christakis, A.; Conneau, A.; Kamali, A.; Jabri, A.; Moyer, A.; Tam, A.; Crookes, A.; Tootoochian, A.; Tootoonchian, A.; Kumar, A.; Vallone, A.; Karpathy, A.; Braunstein, A.; Cann, A.; Codispoti, A.; Galu, A.; Kondrich, A.; Tulloch, A.; Mishchenko, A.; Baek, A.; Jiang, A.; Pelisse, A.; Woodford, A.; Gosalia, A.; Dhar, A.; Pantuliano, A.; Nayak, A.; and Oliver, A. 2024. GPT-4o System Card. arXiv:2410.21276.

Parkinson, M. 1980. The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, 53: 61–65.

Pratapa, A.; Small, K.; and Dreyer, M. 2023. Background Summarization of Event Timelines. In Bouamor, H.; Pino, J.; and Bali, K., eds., *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 8111–8136. Singapore: Association for Computational Linguistics.

Qorib, M. R.; Hu, Q.; and Ng, H. T. 2024. Just What You Desire: Constrained Timeline Summarization with Self-Reflection for Enhanced Relevance. arXiv:2412.17408.

Romanou, A.; Montariol, S.; Paul, D.; Laugier, L.; Aberer, K.; and Bosselut, A. 2023. CRAB: Assessing the Strength of Causal Relationships Between Real-world Events. In Bouamor, H.; Pino, J.; and Bali, K., eds., *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 15198–15216. Singapore: Association for Computational Linguistics.

Saxena, P.; Janzen, S.; and Maass, W. 2024. Newspaper Signaling for Crisis Prediction. In Chang, K.-W.; Lee, A.; and Rajani, N., eds., *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 3: System Demonstrations)*, 166–173. Mexico City, Mexico: Association for Computational Linguistics.

Sun, X.; Liu, M.; and Sima, Z. 2020. A novel cryptocurrency price trend forecasting model based on LightGBM. *Finance Research Letters*, 32: 101084.

Trad, F.; and Chehab, A. 2025. *To Ensemble or Not: Assessing Majority Voting Strategies for Phishing Detection with Large Language Models*, 158–173. Springer Nature Switzerland. ISBN 9783031821509.

Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Ichter, B.; Xia, F.; Chi, E.; Le, Q.; and Zhou, D. 2023. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. arXiv:2201.11903.

Zhang, P.; Xu, K.; and Qi, J. 2023. The impact of regulation on cryptocurrency market volatility in the context of the COVID-19 pandemic—evidence from China. *Economic Analysis and Policy*, 80: 222–246.