Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Leveraging Large Language Models for Sentiment Analysis: Multi-Modal Analysis of Decentraland's MANA Token

Xintong Wu
University of Pennsylvania
Peiting Tsai
Microsoft
Jing Yuan
Microsoft
Michael Yu
Microsoft
Greg Sun
Microsoft
Luyao Zhang
lz183@duke.edu

Duke Kunshan Univeristy https://orcid.org/0000-0002-1183-2254

Research Article

Keywords: Large Language Models, Sentiment Analysis, Multi-Modal Learning, Cryptocurrency Prediction, Metaverse, Blockchain Economics

Posted Date: April 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9317342/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Leveraging Large Language Models for Sentiment Analysis: Multi-Modal Analysis of Decentraland's MANA Token

Xintong Wu
University of Pennsylvania
Philadelphia, PA, USA

Peiting Tsai, Jing Yuan, Michael Yu, and Greg Sun
Microsoft
Beijing, China

Luyao Zhang*
Duke Kunshan University
Kunshan, Jiangsu, China

***Abstract*—Decentraland, a decentralized virtual reality platform operating within the expanding Metaverse ecosystem, utilizes its native MANA token to facilitate virtual asset transactions and governance. This study investigates the integration of Discord community sentiment with multi-modal financial data to enhance cryptocurrency price prediction within virtual world economies. We address: (1) identifying sentiment patterns within Decentraland's Discord community, and (2) evaluating the impact of multi-modal features on token return forecasting. Using a BERT-based large language model for sentiment analysis, we develop two LSTM architectures: a baseline incorporating historical prices and a multi-modal variant integrating sentiment scores, trading volume, and market capitalization. Results indicate predominantly neutral community sentiment with a positive skew. The multi-modal model significantly outperforms the price-only baseline in prediction accuracy. These findings demonstrate the predictive value of community-derived signals for virtual economy forecasting and establish a foundation for future research at the intersection of immersive virtual environments, natural language processing, and cryptocurrency market analysis.**

***Index Terms*—Large Language Models, Sentiment Analysis, Multi-Modal Learning, Cryptocurrency Prediction, Metaverse, Blockchain Economics**

I. Introduction

The emergence of blockchain-enabled immersive virtual environments—the Metaverse—has catalyzed interest in persistent digital economies where users transact virtual assets through native cryptocurrencies [1, 2]. Decentraland represents a paradigmatic case of this convergence, combining virtual reality with decentralized infrastructure to create economic spheres with unique market dynamics [3]. Its native token, MANA, facilitates transactions for virtual goods and land parcels while conferring governance rights through a decentralized autonomous organization (DAO), rendering it ideal for examining sentiment dynamics in nascent virtual economies. Investor sentiment from social media constitutes a significant predictor of market behavior, prompting extensive applications of sentiment analysis in financial forecasting [4]. Within Decentraland, Discord serves as the primary communication channel where user discussions may influence MANA price dynamics. Recent advances in large language models (LLMs) have enhanced sentiment analysis capabilities, with multi-agent frameworks improving financial classification accuracy [5] and specialized models advancing emotion detection [6].

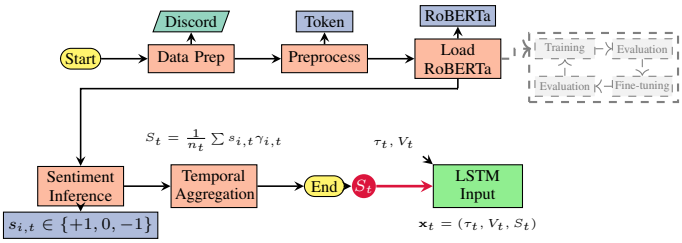


Fig. 1: LLM-based sentiment analysis pipeline. *Red circle highlights sentiment output S_t feeding into downstream LSTM alongside conventional technical indicators such as price τ_t and volume V_t.*

This study constructs a multi-modal dataset integrating historical price data with Discord-derived sentiment, trading volume, and market capitalization to forecast token returns. We investigate two research questions (RQs):

- **(RQ1)** What are the prevailing sentiment patterns among Decentraland Discord users, and to what extent do these reflect platform dynamics?
- **(RQ2)** How does integrating multi-modal features affect token return prediction accuracy relative to price-only models?

Employing a BERT-based architecture for sentiment extraction and LSTM networks for time-series prediction, we evaluate the incremental value of community sentiment for cryptocurrency forecasting. Figure 1 illustrates the LLM-based sentiment analysis pipeline.[1]

*Corresponding author: Luyao Zhang (lz183@duke.edu), Digital Innovation Research Center and Social Science Division, Duke Kunshan University. Address: Duke Avenue No.8, Kunshan, Suzhou, Jiangsu, China, 215316.
Acknowledgments: Luyao Zhang acknowledges support from the National Science Foundation of China (NSFC) for the project titled "Trust Mechanism Design on Blockchain: An Interdisciplinary Approach of Game Theory, Reinforcement Learning, and Human-AI Interactions" (Grant No. 12201266). Xintong Wu is grateful for the support provided by the Office of Academic Services, the Summer Research Scholar Program, and the Student Experiential Learning Fellow (SELF) Program at Duke Kunshan University, supervised by Prof. Luyao Zhang. Xintong also extends appreciation to all members of the One-Person Entrepreneur (OPE) group at Microsoft China for their valuable inspiration.

[1]All data and code are publicly accessible at: https://github.com/Xintong1122/Decentraland_Senti

II. Data

A. Data Collection

1) Numerical Data: Daily OHLCV data for MANA (ERC-20 utility token [7–9]) spanning $t \in [1, T]$ where $T = 366$ days (June 15, 2023–June 14, 2024) comprise open, high, low, close prices ($P_t^O, P_t^H, P_t^L, P_t^C$), trading volume V_t, and market capitalization M_t (USD), obtained from *CoinMarketCap*.[2]

2) Textual Data: Discord *#general* channel messages $\mathcal{M} = \{m_i\}_{i=1}^n$ ($n = 5,883$) were harvested via *DiscordChatExporter*,[3] yielding timestamps t_i, anonymized author IDs α_i, content c_i, and engagement metrics. Post-filtering invalid entries, $n' = 5,513$ messages ($\mathcal{M}' \subset \mathcal{M}$) comprised the final corpus.

B. Data Processing

1) Price Metric: The typical price $\tau_t = \frac{1}{3}(P_t^H + P_t^L + P_t^C)$ serves as the daily indicator, smoothing intraday fluctuations following standard technical analysis practice [10].

2) Return Calculation: Continuous returns $r_t = \ln(\tau_{t+1}/\tau_t) = \ln(\tau_{t+1}) - \ln(\tau_t)$ stabilize variance and ensure temporal comparability for ERC-20 token markets [11].

III. Methodology

A. Sentiment Analysis

1) LLM Selection: We employ the RoBERTa-based classifier *cardiffnlp/twitter-roberta-base-sentiment-latest*[4] from Hugging Face, pretrained on social media corpora for informal text classification [12, 13].

2) Aggregation: Each message i at time t maps to sentiment $s_{i,t} \in \{+1, 0, -1\}$ (positive, neutral, negative) with confidence $\gamma_{i,t}$. Daily sentiment aggregates as $S_t = \frac{1}{n_t}\sum_{i=1}^{n_t} s_{i,t}\gamma_{i,t}$, where n_t denotes daily message volume [14]. Discretization thresholds define $S_t \in [-1, -0.1)$ as negative, $[-0.1, 0.1]$ as neutral, and $(0.1, 1]$ as positive.

B. Return Prediction

1) LSTM Architecture: Long Short-Term Memory networks model sequential dependencies $\{(\mathbf{x}_t, r_t)\}_{t=1}^T$ via gated mechanisms [15–17]:

$$i_t = \sigma(W_i \cdot [h_{t-1}, \mathbf{x}_t] + b_i), \quad o_t = \sigma(W_o \cdot [h_{t-1}, \mathbf{x}_t] + b_o) \quad (1)$$

where i_t, o_t denote input and output gates, σ the sigmoid function, and W, b learned parameters. Model specifications:

- **Baseline:** $\mathbf{x}_t = \tau_t$ (scalar, typical price only)
- **Multi-modal:** $\mathbf{x}_t = (\tau_t, V_t, M_t, S_t)' \in \mathcal{R}^4$

2) Evaluation Metrics: Predictive performance is quantified via MSE $= T^{-1}\sum_{t=1}^T (r_t - \hat{r}_t)^2$, MAE $= T^{-1}\sum_{t=1}^T |r_t - \hat{r}_t|$, and $R^2 = 1 - \sum_{t=1}^T (r_t - \hat{r}_t)^2 / \sum_{t=1}^T (r_t - \bar{r})^2$ [18–22].

[2]https://coinmarketcap.com/

[3]https://github.com/Tyrrrz/DiscordChatExporter

[4]https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest

IV. Results

Figure 2 presents the dual analysis of community sentiment dynamics and predictive model performance. The bar plot illustrates the distribution of Discord sentiment over the 366-day observation period: 26.94% positive ($s_{i,t} = +1$), 60.96% neutral ($s_{i,t} = 0$), and 12.10% negative ($s_{i,t} = -1$). This predominance of neutral sentiment with a positive skew indicates stable community affect, with daily aggregate scores S_t concentrating within $[0, 1]$. Notably, negative sentiment spikes correlate primarily with user-level technical issues rather than exogenous market events [23, 24], suggesting these data reflect platform user experience rather than macroeconomic signals.

The time-series plots compare the baseline LSTM (price-only feature $\mathbf{x}_t = \tau_t$) against the multi-modal architecture ($\mathbf{x}_t = (\tau_t, V_t, M_t, S_t)' \in \mathcal{R}^4$). The multi-modal model improves prediction accuracy, reducing MSE from 0.002100 to 0.001528 and MAE from 0.0297 to 0.0241. However, both models exhibit negative R^2 values (baseline: -0.4185; multi-modal: -0.0321), indicating limited explanatory power for directional trends. While auxiliary features enhance short-term precision, they inadequately capture broader market dynamics, consistent with [25].

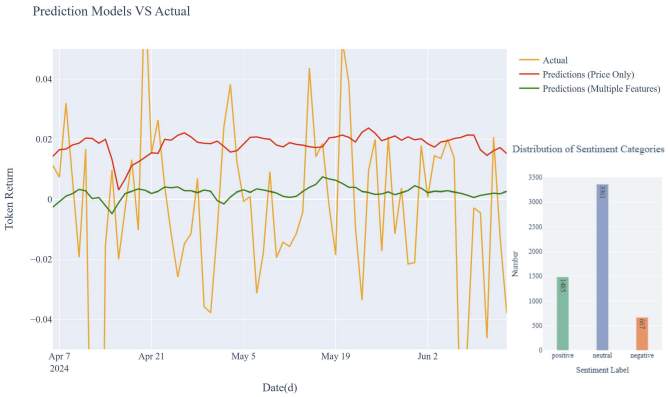


Fig. 2: Price Prediction and Sentiment Categories.

V. Conclusion

Our study contributes to metaverse and blockchain-empowered virtual economy research [4, 26–32] and the literature on user-generated content and community sentiment [33–36]. By applying large language models to financial sentiment [37–45], we extend methodological approaches in sentiment analysis [46–52] to the cryptocurrency domain [53–59]. Results show that incorporating sentiment enhances price-only models [60–62], though gains remain limited as Discord sentiment reflects platform-specific user experience rather than macro-market conditions. Future work should integrate additional data sources [63], improve domain-specific modeling, and consider scalability and security aspects [1].[5]

[5]Detailed contributions to each literature stream are provided in Appendix A.

References

[1] H. Huang, Z. Yin, Q. Yang, T. Li, X. Luo, L. Zhou, and Z. Zheng, "Scalability and security of blockchain-empowered metaverse: A survey," *IEEE Open Journal of the Computer Society*, 2024.

[2] V. T. Truong, L. Le, and D. Niyato, "Blockchain meets metaverse and digital asset management: A comprehensive survey," *Ieee Access*, vol. 11, pp. 26 258–26 288, 2023.

[3] C. Goanta, "Selling land in decentraland: The regime of non-fungible tokens on the ethereum blockchain under the digital content directive," *Disruptive technology, legal innovation, and the future of real estate*, pp. 139–154, 2020.

[4] B. Guidi and A. Michienzi, "Social games and blockchain: exploring the metaverse of decentraland," in *2022 IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW)*. IEEE, 2022, pp. 199–204.

[5] F. Xing, "Designing heterogeneous llm agents for financial sentiment analysis," *ACM Trans. Manage. Inf. Syst.*, Aug. 2024, just Accepted. [Online]. Available: https://doi.org/10.1145/3688399

[6] Z. Liu, K. Yang, Q. Xie, T. Zhang, and S. Ananiadou, "Emollms: A series of emotional large language models and annotation tools for comprehensive affective analysis," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, ser. KDD '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 5487–5496. [Online]. Available: https://doi.org/10.1145/3637528.3671552

[7] Decentraland Foundation, "Decentraland white paper 2.0," Decentraland Foundation, Tech. Rep., 2024, accessed: 2024. [Online]. Available: https://decentraland.org/whitepaper2.pdf

[8] E. Ordano, A. Meilich, Y. Jardi, and M. Araoz, "Decentraland: A blockchain-based virtual world," Decentraland, Tech. Rep., 2017. [Online]. Available: https://decentraland.org/whitepaper.pdf

[9] Y. Hu, Z. Li, J. Chen, Y. Gao, and Y. Liu, "Exploring the decentraland economy: Multifaceted parcel attributes, key insights, and benchmarking," *arXiv preprint arXiv:2404.07533*, 2025. [Online]. Available: https://arxiv.org/abs/2404.07533

[10] J. J. Murphy, *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin, 1999.

[11] H. T. Heinonen, A. Semenov, and V. Boginski, "Collective behavior of price changes of erc-20 tokens," in *Computational Data and Social Networks: 9th International Conference, CSoNet 2020, Dallas, TX, USA, December 11–13, 2020, Proceedings 9*. Springer, 2020, pp. 487–498.

[12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, 2019, pp. 4171–4186. [Online]. Available: https://aclanthology.org/N19-1423/

[13] M. Masala, S. Ruseti, and M. Dascalu, "Robert–a romanian bert model," in *Proceedings of the 28th International Conference on Computational Linguistics*, 2020, pp. 6626–6637.

[14] B. Zhang, H. Yang, T. Zhou, M. Ali Babar, and X.-Y. Liu, "Enhancing financial sentiment analysis via retrieval augmented large language models," in *Proceedings of the fourth ACM international conference on AI in finance*, 2023, pp. 349–356.

[15] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," *Neural computation*, vol. 31, no. 7, pp. 1235–1270, 2019.

[16] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[17] R. C. Staudemeyer and E. R. Morris, "Understanding lstm–a tutorial into long short-term memory recurrent neural networks," *arXiv preprint arXiv:1909.09586*, 2019.

[18] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation," *Peerj computer science*, vol. 7, p. e623, 2021.

[19] O. Köksoy, "Multiresponse robust design: Mean square error (mse) criterion," *Applied Mathematics and Computation*, vol. 175, no. 2, pp. 1716–1729, 2006.

[20] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance," *Climate research*, vol. 30, no. 1, pp. 79–82, 2005.

[21] T. Chai and R. R. Draxler, "Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature," *Geoscientific model development*, vol. 7, no. 3, pp. 1247–1250, 2014.

[22] A. C. Cameron and F. A. Windmeijer, "An r-squared measure of goodness of fit for some common nonlinear regression models," *Journal of econometrics*, vol. 77, no. 2, pp. 329–342, 1997.

[23] D. Foundation, "Decentraland's weekly newsletter - jan 4," 2023. [Online]. Available: https://decentraland.beehiiv.com/p/weekly-newsletter-january-4

[24] L. Gibbons, "Decentraland art week 2024 kicks off," Mar 2024. [Online]. Available: https://www.blockleaders.io/news/decentraland-art-week-2024-kicks-off

[25] T.-G. Lee, "A test method for the convergence of the metaverse and blockchain," in *2024 26th International Conference on Advanced Communications Technology (ICACT)*. IEEE, 2024, pp. 321–326.

[26] H. Duan, J. Li, S. Fan, Z. Lin, X. Wu, and W. Cai, "Metaverse for social good: A university campus prototype," in

Proceedings of the 29th ACM international conference on multimedia, 2021, pp. 153–161.

[27] H. D. Le, V. T. Truong, and L. B. Le, "Blockchain-empowered metaverse: Decentralized crowdsourcing and marketplace for trading machine learning data and models," *IEEE Access*, 2024.

[28] B. Wu and B. Wu, "Metaverse: The world reimagined," in *Blockchain for teens: With case studies and examples of blockchain across various industries*. Springer, 2022, pp. 267–313.

[29] E. Den Yeoh, T. Chung, and Y. Wang, "Metaverse in investment using sentiment analysis and machine learning," in *Strategies and Opportunities for Technology in the Metaverse World*. IGI Global, 2023, pp. 78–113.

[30] L. Zhang, Y. Quan, J. Cao, K. Z. Zhou, and X. Tong, "Leveraging social media sentiments and ethical signals for NFT valuation," in *2024 IEEE 24th International Conference on Software Quality, Reliability, and Security Companion (QRS-C)*. IEEE, 2024, pp. 206–215. [Online]. Available: https://ieeexplore.ieee.org/document/10727085

[31] T. Yan, S. Li, B. Kramer, L. Zhang, and C. J. Tessone, "A data engineering framework for ethereum beacon chain rewards: From data collection to decentralization metrics," *Scientific Data*, vol. 12, no. 1, p. 519, 2025. [Online]. Available: https://www.nature.com/articles/s41597-025-04623-7

[32] N. Chemaya, L. W. Cong, E. Joergensen, D. Liu, and L. Zhang, "A dataset of uniswap daily transaction indices by network," *Scientific Data*, vol. 12, no. 1, p. 93, 2025. [Online]. Available: https://www.nature.com/articles/s41597-024-04042-0

[33] H. J. Cheong and M. A. Morrison, "Consumers' reliance on product information and recommendations found in ugc," *Journal of interactive advertising*, vol. 8, no. 2, pp. 38–49, 2008.

[34] A. Z. Bahtar and M. Muda, "The impact of user–generated content (ugc) on product reviews towards online purchasing–a conceptual framework," *Procedia Economics and Finance*, vol. 37, pp. 337–342, 2016.

[35] A. Micu, A. E. Micu, M. Geru, and R. C. Lixandroiu, "Analyzing user sentiment in social media: Implications for online marketing strategy," *Psychology & Marketing*, vol. 34, no. 12, pp. 1094–1100, 2017.

[36] S. Y. Yang, S. Y. K. Mo, and A. Liu, "Twitter financial community sentiment and its predictive relationship to stock market movement," *Quantitative Finance*, vol. 15, no. 10, pp. 1637–1656, 2015.

[37] E. Kasneci, K. Seßler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, G. Groh, S. Günnemann, E. Hüllermeier *et al.*, "Chatgpt for good? on opportunities and challenges of large language models for education," *Learning and individual differences*, vol. 103, p. 102274, 2023.

[38] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang *et al.*, "A survey on evaluation of large language models," *ACM Transactions on Intelligent Systems and Technology*, vol. 15, no. 3, pp. 1–45, 2024.

[39] H. Huang, O. Zheng, D. Wang, J. Yin, Z. Wang, S. Ding, H. Yin, C. Xu, R. Yang, Q. Zheng *et al.*, "Chatgpt for shaping the future of dentistry: the potential of multi-modal large language model," *International Journal of Oral Science*, vol. 15, no. 1, p. 29, 2023.

[40] H. Li, "Deep learning for natural language processing: advantages and challenges," *National Science Review*, vol. 5, no. 1, pp. 24–26, 2018.

[41] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "Roberta: A robustly optimized bert pretraining approach," *arXiv preprint arXiv:1907.11692*, 2019.

[42] H. Xu, B. Liu, L. Shu, and P. S. Yu, "Bert post-training for review reading comprehension and aspect-based sentiment analysis," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, 2019, pp. 2324–2335. [Online]. Available: https://aclanthology.org/N19-1242/

[43] T. Zhan, C. Shi, Y. Shi, H. Li, and Y. Lin, "Optimization techniques for sentiment analysis based on llm (gpt-3)," *arXiv preprint arXiv:2405.09770*, 2024.

[44] P. Törnberg, D. Valeeva, J. Uitermark, and C. Bail, "Simulating social media using large language models to evaluate alternative news feed algorithms," *arXiv preprint arXiv:2310.05984*, 2023.

[45] X. Deng, V. Bashlovkina, F. Han, S. Baumgartner, and M. Bendersky, "What do llms know about financial markets? a case study on reddit market sentiment analysis," in *Companion Proceedings of the ACM Web Conference 2023*, 2023, pp. 107–110.

[46] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams engineering journal*, vol. 5, no. 4, pp. 1093–1113, 2014.

[47] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731–5780, 2022.

[48] D. M. E.-D. M. Hussein, "A survey on sentiment analysis challenges," *Journal of King Saud University-Engineering Sciences*, vol. 30, no. 4, pp. 330–338, 2018.

[49] M. D. Devika, C. Sunitha, and A. Ganesh, "Sentiment analysis: a comparative study on different approaches," *Procedia Computer Science*, vol. 87, pp. 44–49, 2016.

[50] P. Gonçalves, M. Araújo, F. Benevenuto, and M. Cha, "Comparing and combining sentiment analysis methods," in *Proceedings of the first ACM conference on Online social networks*, 2013, pp. 27–38.

[51] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowledge and Information Systems*, vol. 60, pp. 617–663, 2019.

[52] F. Neri, C. Aliprandi, F. Capeci, and M. Cuadros, "Sentiment analysis on social media," in *2012 IEEE/ACM international conference on advances in social networks analysis and mining*. IEEE, 2012, pp. 919–926.
[53] A. ElBahrawy, L. Alessandretti, A. Kandler, R. Pastor-Satorras, and A. Baronchelli, "Evolutionary dynamics of the cryptocurrency market," *Royal Society open science*, vol. 4, no. 11, p. 170623, 2017.
[54] B. M. Henrique, V. A. Sobreiro, and H. Kimura, "Literature review: Machine learning techniques applied to financial market prediction," *Expert Systems with Applications*, vol. 124, pp. 226–251, 2019.
[55] C. Y.-H. Chen and C. M. Hafner, "Sentiment-induced bubbles in the cryptocurrency market," *Journal of Risk and Financial Management*, vol. 12, no. 2, p. 53, 2019.
[56] D. Zimbra, A. Abbasi, D. Zeng, and H. Chen, "The state-of-the-art in twitter sentiment analysis: A review and benchmark evaluation," *ACM Transactions on Management Information Systems (TMIS)*, vol. 9, no. 2, pp. 1–29, 2018.
[57] C. A. Melton, O. A. Olusanya, N. Ammar, and A. Shaban-Nejad, "Public sentiment analysis and topic modeling regarding covid-19 vaccines on the reddit social media platform: A call to action for strengthening vaccine confidence," *Journal of Infection and Public Health*, vol. 14, no. 10, pp. 1505–1512, 2021.
[58] A. S. M. Alharbi and E. de Doncker, "Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information," *Cognitive Systems Research*, vol. 54, pp. 50–61, 2019.
[59] Y. Quan, X. Wu, W. Deng, and L. Zhang, "Decoding social sentiment in dao: A comparative analysis of blockchain governance communities," in *2024 IEEE 24th International Conference on Software Quality, Reliability, and Security Companion (QRS-C)*. IEEE, 2024, pp. 216–224. [Online]. Available: https://doi.org/10.1109/QRS-C63300.2024.00037
[60] S. Siami-Namini and A. S. Namin, "Forecasting economics and financial time series: Arima vs. lstm," *arXiv preprint arXiv:1803.06386*, 2018.
[61] S. Athey and G. W. Imbens, "Machine learning methods that economists should know about," *Annual Review of Economics*, vol. 11, no. 1, pp. 685–725, 2019.
[62] J. D. Hamilton, *Time series analysis*. Princeton university press, 2020.
[63] Q. Li, J. Tan, J. Wang, and H. Chen, "A multimodal event-driven lstm model for stock prediction using online news," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 10, pp. 3323–3337, 2020.

APPENDIX A
DETAILED CONTRIBUTIONS TO RELATED WORK

This study advances four interconnected research streams.

Metaverse and Decentraland Research. While prior work has established the technological foundations of the Metaverse [26] and Decentraland's specific implementation [4, 29], this study is among the first to empirically link platform-specific community sentiment to token price dynamics (see Table III for performance metrics). We demonstrate that DAO-governed virtual worlds [28] generate user-generated content [33, 34] with measurable but bounded predictive signal for MANA returns, extending research on blockchain-empowered metaverse economies [27] and the importance of monitoring user sentiment for platform governance [35, 36]. The daily sentiment trajectory shown in Figure 3 illustrates the temporal dynamics of community engagement.

Sentiment Analysis Methodology. Our work bridges lexicon-based and machine-learning approaches [46, 49, 50] with modern LLM applications. By validating that general-domain sentiment surveys [47, 48, 51, 52] require domain adaptation for metaverse contexts, we contribute to methodological refinements in social media sentiment analysis [56–58] and crypto-specific opinion mining [55, 59]. The textual data structure summarized in Table II enables these analyses.

LLM Applications in Financial Forecasting. We advance the deployment of large language models [37–40] for specialized financial sentiment tasks. Utilizing architectures like BERT and RoBERTa [41–43], we test their efficacy on Discord data, contributing to research on LLM-based social media simulation [44] and Reddit-style market sentiment analysis [45] within decentralized virtual communities.

Cryptocurrency Prediction and Volatility. Addressing the evolutionary dynamics of crypto markets [53], our LSTM-based multi-modal framework responds to calls for advanced prediction methodologies [54]. We contribute to time series forecasting literature [60, 62] by quantifying the marginal value of sentiment features beyond technical indicators [61], demonstrating both the potential and limitations of community-specific data for crypto-volatility modeling.

APPENDIX B
DATA SPECIFICATIONS AND EMPIRICAL RESULTS

We detail the dataset structures and empirical benchmarks supporting the multi-modal analysis.

A. Market Data Structure

Table I defines the quantitative financial variables extracted from Decentraland's MANA token markets. These time-series features constitute the baseline price-only model inputs.

TABLE I: Numerical Data: MANA Price, Volume, Market Cap

Variable	Description	Unit
Date	Trading timestamp	—
Open Price	Opening market price	$ USD
High Price	Daily maximum price	$ USD
Low Price	Daily minimum price	$ USD
Close Price	Closing market price	$ USD
Volume	Trading volume (tokens exchanged)	MANA
Market Cap	Total market capitalization	$ USD

Note: All price and market cap values are denominated in US Dollars.

B. Sentiment Corpus Structure

Table II outlines the Discord message schema. The sentiment labels and scores are derived via LLM-based classification of the Content field.

TABLE II: Textual Data: Users' Messages (Discord Corpus)

Field	Description	Type
Date & Time	Message timestamp (UTC)	
Author ID	Anonymized user identifier	
Content	Raw message text body	
Attachments	Media/files shared	
Reactions	Emoji responses count	
Label	Assigned sentiment category	
Score	Numerical sentiment polarity (-1 to $+1$)	

Note: Sentiment labels and scores derived via LLM-based classification.

C. Model Performance Benchmarks

Table III presents the eight-fold cross-validation results comparing price-only LSTM baselines against multi-modal variants (Price + Sentiment). The average improvements indicate that sentiment features provide marginal but consistent predictive enhancement, though negative R^2 values in both configurations suggest substantial unmodeled volatility drivers.

D. Temporal Visualization

Figure 3 displays the daily averaged sentiment scores derived from the Discord corpus (Table II), revealing temporal fluctuations in community mood. Figure 4 overlays predicted returns from the multi-modal model against realized returns, visualizing the systematic underestimation of volatility discussed.

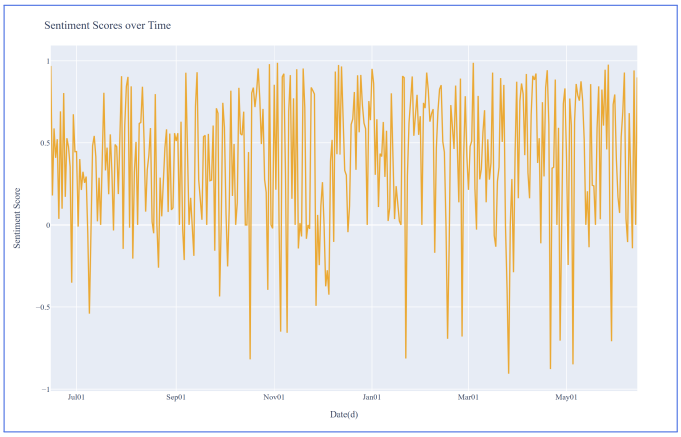


Fig. 3: Daily Average Sentiment Score (Decentraland Discord Community).

TABLE III: Comparison of Prediction Models (8 Independent Runs)

Run	Model	MSE ($\times 10^{-4}$)	MAE	R^2
1	Price Only	5.50	0.01835	−0.192
	+Sentiment	5.53↑	0.01816↓	−0.198↓
2	Price Only	5.30	0.01840	−0.148
	+Sentiment	5.27↓	0.01949↑	−0.141↑
3	Price Only	5.79	0.01853	−0.256
	+Sentiment	5.57↓	0.01784↓	−0.207↑
4	Price Only	5.79	0.01853	−0.256
	+Sentiment	5.57↓	0.01784↓	−0.207↑
5	Price Only	6.75	0.01981	−0.463
	+Sentiment	5.26↓	0.01916↓	−0.140↑
6	Price Only	5.96	0.01870	−0.291
	+Sentiment	5.27↓	0.01807↓	−0.143↑
7	Price Only	6.34	0.01913	−0.375
	+Sentiment	5.46↓	0.01932↑	−0.184↑
8	Price Only	5.31	0.01818	−0.150
	+Sentiment	5.24↓	0.01925↑	−0.136↑
Avg	**Price Only**	**5.84**	**0.0187**	**−0.266**
	+Sentiment	**5.40↓**	**0.0186↓**	**−0.170↑**

Note: MSE scaled by 10^4 for readability. Legend: ↓=Lower error/higher R^2 (better); ↑=Worse.

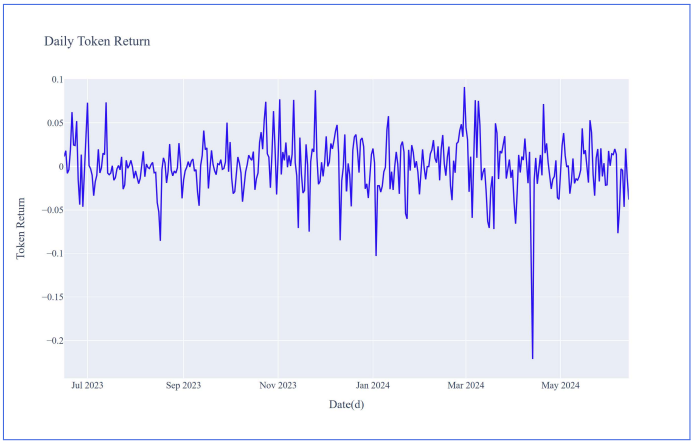


Fig. 4: Daily Token Return.