



Quantitative Assessment of AI-Driven Sentiment Classification and Price Prediction in Cryptocurrency Markets

Khuwaish Chulani and Dr. Danling Jiang
Blockchain Business Lab



ABSTRACT

This project evaluates the accuracy of AI-generated sentiment predictions in forecasting short-term cryptocurrency price movements. Using daily cryptocurrency data and model-predicted sentiment labels, the analysis compares predicted trends against actual outcomes. A custom evaluation framework measured precision, recall, and F1 score across bearish, neutral, and bullish classes. Results show stronger performance in bearish predictions and inconsistencies in bullish forecasting, highlighting both the potential and limitations of sentiment-driven models.

INTRODUCTION

This project investigates the use of AI to forecast short-term cryptocurrency price behavior based on daily market data. A large language model (LLM) was prompted with structured data—including price, volume, market capitalization, and percent change metrics—for the top 100 cryptocurrencies each day. The model generated next-day predictions for price and market cap, which were then compared with actual values to infer directional price movement. Based on this, sentiment was classified into three categories: bearish (-1), neutral (0), or bullish (1). These AI-labeled sentiments were then evaluated against real market outcomes using precision, recall, and F1 score metrics. This work aims to assess the reliability and limitations of AI-driven sentiment predictions in dynamic crypto market environments.

METHODS

- A robust data architecture was developed using Python, integrating the CoinMarketCap API for market data collection and the Perplexity AI API for prediction generation. The system was supported by pandas, json, matplotlib, and seaborn for processing, evaluation, and visualization.
- Daily market data for the top 100 cryptocurrencies—including price, volume, market capitalization, and percent change metrics—was collected throughout March 2025 and stored in structured JSON format.
- A large language model (LLM) was queried daily in a simulated real-time environment to generate next-day predictions for price and market cap, without access to future data.
- Sentiment classification was derived by comparing predicted and actual price movement: Bullish (1) for upward deviation, Neutral (0) for price stability within $\pm 0.1\%$, and Bearish (-1) for downward deviation. Ground truth sentiment was calculated using the same logic.
- A custom evaluation pipeline was developed to compute performance metrics—precision, recall, and F1 score—for each sentiment class.
- Visualization outputs included bar charts, time series plots, accuracy heatmaps, and correlation matrices linking sentiment to market metrics.

RESULTS

Sentiment Class	Precision	Recall	F1 Score	Accuracy	Support
-1	0.2679	0.832	0.4053	0.832	732
0	0.5416	0.2228	0.3157	0.2228	1315
1	0.4038	0.0683	0.1168	0.0683	923

Figure 1. Sentiment Class Performance Metrics

This table presents precision, recall, F1 score, accuracy, and support for each predicted sentiment class: bearish (-1), neutral (0), and bullish (1). The model shows strong recall for bearish predictions but lower precision, indicating it captures most bearish cases but with more false positives. Neutral sentiment achieves the highest precision, while bullish predictions display the weakest overall performance across all metrics. These results highlight class-dependent variance in prediction reliability.

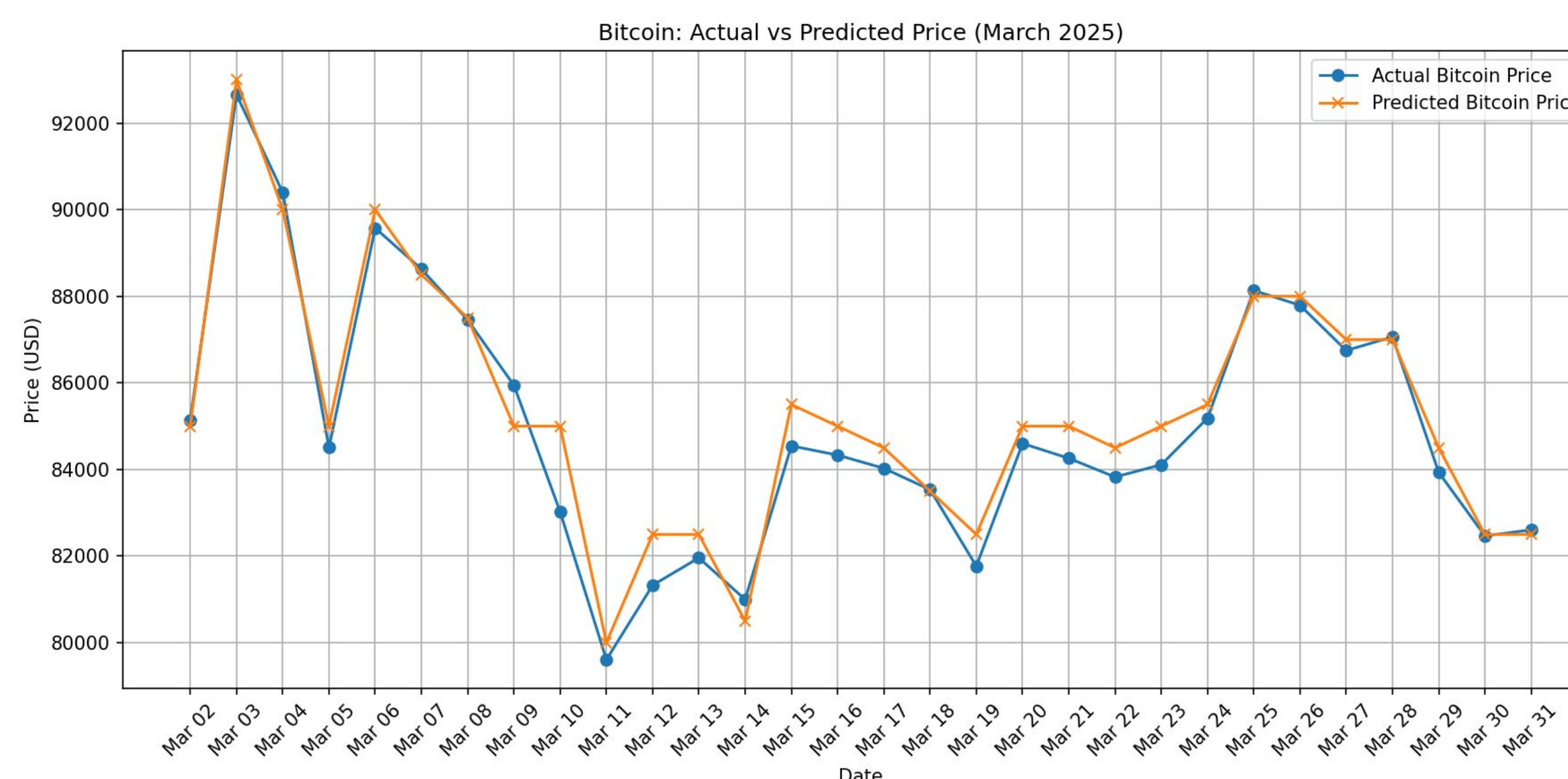


Figure 2. Actual vs Predicted Bitcoin Prices (March 2025)

This time series chart compares actual daily Bitcoin prices with AI-generated next-day predictions throughout March 2025. The predicted values closely follow the overall trend of real prices, demonstrating the model's ability to capture short-term market direction. While some deviations are present—especially during periods of volatility—the model aligns well at key turning points and shows minor average error. These results suggest the AI predictions are reasonably accurate for trend tracking, though less precise in exact price magnitude.

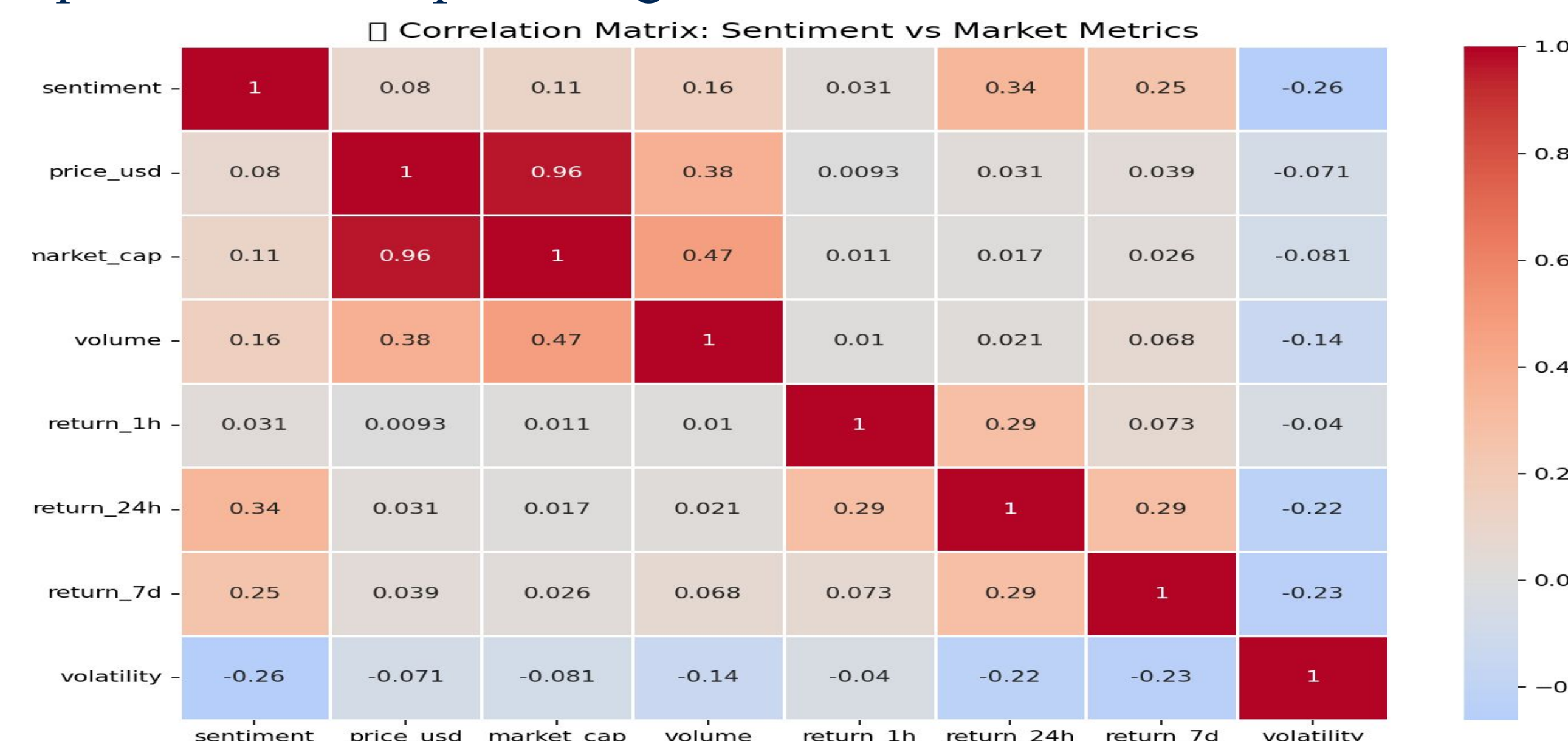


Figure 3. Correlation Matrix: Sentiment vs Market Metrics

This heatmap displays Pearson correlations between AI-predicted sentiment and key market metrics. Sentiment aligns most with 24-hour returns ($r = 0.34$) and 7-day returns ($r = 0.25$), indicating predictive relevance for short-term movement. Moderate correlation with volume ($r = 0.16$) and a negative correlation with volatility ($r = -0.26$) suggest reduced confidence in unstable markets. Price and market cap show weak correlations, emphasizing the model's focus on directional trends.

DISCUSSION & CONCLUSION

- The findings suggest that AI-generated sentiment can partially anticipate short-term cryptocurrency trends. While the model achieved high recall in detecting downward trends, it showed reduced accuracy when predicting bullish sentiment.
- Overall, the predicted prices—especially for leading coins like Bitcoin—closely mirrored actual market values, reinforcing the capability of LLMs to simulate realistic next-day forecasts using only current-day data.
- This study demonstrates the promise of combining structured market data with AI-driven forecasting to assess directional movement across diverse assets. However, the results also emphasize the need for improved modeling techniques and greater robustness to market volatility in order to boost real-world applicability.

IMPLICATIONS

- This research highlights the potential for integrating AI-generated sentiment into short-term crypto forecasting, offering a foundation for developing sentiment-aware trading tools and data-driven decision systems.
- The varied performance across sentiment classes suggests future models could be refined or ensemble methods explored to enhance prediction consistency and reduce false positives.
- Findings may be useful for analysts, developers, or institutions exploring AI applications in financial markets—particularly in the context of directional movement rather than price magnitude.
- Given market volatility, further research could explore hybrid models that combine sentiment, technical indicators, and real-time data streams to better capture crypto market behavior.

KEY REFERENCES

- <https://coinmarketcap.com/>

ACKNOWLEDGEMENT

- Thank you to Dr. Jiang, Prattusha Biswas, and the BBL for the invaluable support and guidance throughout my research endeavour