Algorithmic Trading in Cryptocurrency Markets

Vasilis Michalakis
The University of Adelaide
a1887068@adelaide.edu.au

Jinan Zou
The University of Adelaide
Australian Institute for Machine Learning
jinan.zou@adelaide.edu.au

Haiyao Cao
The University of Adelaide
Australian Institute for Machine Learning
haiyao.cao@adelaide.edu.au

ABSTRACT

Algorithmic trading refers to the automated execution of buy and sell orders in financial markets. Such algorithms may leverage machine learning techniques or follow rule-based strategies, including arbitrage, technical analysis, and momentum trading. However, in the retail context, profitable trading of cryptocurrency remains constrained by data sparsity, overfitting, and volatile market conditions. We propose a hybrid strategy integrating sentiment analysis of social media and news sources with technical analysis. Our method outperforms both a passive portfolio and a time-series momentum strategy, but falls short of the benchmark sentiment strategy.

1 INTRODUCTION

Cryptocurrency has emerged as a significant alternative asset, due to its growing popularity and unique characteristics [1]. Broadly, cryptocurrencies can be categorized into three types of "coins": Bitcoin, the original cryptocurrency with the largest market capitalisation; altcoins, which encompass all cryptocurrencies other than Bitcoin; and meme coins, which are typically named after internet memes or pop culture phenomena [2]. Cryptocurrency markets also feature stablecoins, which are cryptocurrencies pegged to a fiat currency, commodity, or financial instrument; for example, Tether (USDT), a widely traded stablecoin pegged to the US Dollar (USD) [3]. Furthermore, Decentralised Exchanges (DEXs) are available as an alternative trading venue, offering blockchain-based order settlement [4].

Trading strategies generally fall into two categories: fundamental analysis and technical analysis [5]. Fundamental analysis evaluates an asset's intrinsic value to determine whether it is undervalued or overvalued [5]. In the context of cryptocurrency, this involves examining factors such as the project's vision, the development team, whitepaper, engagement of the technical community, token supply (fixed or expanding), and on-chain metrics [5]. Technical analysis, however, focuses on patterns, using historical data to forecast future price movements [5]. The price of a cryptocurrency is dependent on a variety of blockchain, market, macro-economic and political factors [6] (see A.1.1).

The Efficient Market Hypothesis (EMH) suggests that prices reflect all available information, making it impossible to consistently outperform passive investment strategies [7]. In contrast, behavioural finance suggests that price deviations from intrinsic value arise due to cognitive biases such as over-optimism, overreaction, and bounded rationality [8], providing the framework that underpins our approach.

A typical quantitative trading system consists of signal generation, position control, and asset management components [9]. These systems may use machine learning techniques or rule-based strategies, including arbitrage, technical analysis, and momentum trading. As retail traders, we propose and back-test a hybrid strategy that integrates traditional technical indicators with DeepSeek-powered sentiment analysis of social media and news sources. This strategy is applied to a diversified portfolio, consisting of Bitcoin, altcoins, and meme coins.

1.1 Motivation

The primary objective of any active asset management strategy is to outperform a passive strategy (such as a representative index). Research into algorithmic trading is motivated by its potential to enhance speed, precision, and rationality in financial markets, utilizing computational power to process vast amounts of data in real-time without the limitations of human emotions, error or fatigue [10]. However, research in this field is inherently limited, as papers outlining profitable strategies are unlikely to be published due to the incentive to capitalize on them privately. Furthermore, the tendency of strategies to overfit, combined with the sparse and unstructured nature of available data, inhibit real-world performance [11, 12], necessitating further research.

Sentiment-based trading research is motivated by the speculative nature of cryptocurrency markets [11], especially meme coins [2]. Existing literature often treats sentiment data as equally valuable, overlooking engagement levels and author credibility, which can improve the predictive quality of sentiment signals [13]. Additionally, research integrating machine-learning models with traditional technical analysis rules remains limited [14]. To address these gaps, our study combines engagement-weighted sentiment with technical analysis to better capture market behaviour and trends.

2 LITERATURE REVIEW

2.1 Sentiment Analysis Strategies

Belcastro et al. developed an algorithmic trading strategy using a Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) to predict the price movements of various cryptocurrencies [15]. Their dataset consisted of hourly price and volume data for 38 cryptocurrencies in USDT, sentiment scores of X (formerly Twitter) posts, correlation coefficients between social media posts and price activity, and market data from three closely related cryptocurrencies [15]. They evaluated their algorithm using a stop-loss of 8%, take-profit of 12%, 12-hour prediction to execution lag, and a 3-day lookback window [15]. Their algorithm yielded a mean annual return of 117.40%, back-tested in 2021, across all cryptocurrencies, net of transaction fees. However, their model could be improved by incorporating additional technical indicators to enhance prediction accuracy [16], using dropout layers to reduce the risk of overfitting [17] and validating performance in live markets.

Alternatively, Moradi-Kamali et al. proposed a sentiment-based Bitcoin trading strategy using "Market-Driven Labelling", where sentiment labels are derived from market outcomes [18]. They fine-tuned CryptoBERT with 5-fold cross-validation across three models: Context-Aware (CA), incorporating prior-period features; Temporal Context-Aware (TCA), adding timestamp data; and Context-Unaware (CU), using raw text data from X [18]. Combining daily OHLCV (Open, High, Low, Close, Volume) data, posts from X, and dates of major events, their method generates tweet-level predictions by taking the majority class across the 5 folds. These tweet-level predictions are then aggregated daily by mean or majority vote [18] (refer to A.2.1). Their labelling-based strategy sets stop-loss and take-profit levels based on historical volatility [18]. Back-testing of this method showed a best-case daily return of 7.86% and a Sharpe ratio of 5.07 during a 2017 bull market, using majority sentiment aggregation [18]. However, it remains unclear if labelling parameters were optimized on prior periods, raising concerns of look-ahead bias. Furthermore, the approach remains untested in live markets.

Alternatively, Han et al. proposed a Large Language Model (LLM)-based Multi-Agent System (MAS) to trade Bitcoin, Ethereum, Dogecoin, and Solana across 5-minute, 15-minute, 1-hour, and 4-hour timeframes [19]. Their method, based on CryptoTrade, involves four collaborative LLM agents: a Market Analyst, which receives prompts containing on-chain indicators and technical analysis data; a News Analyst, which processes financial news and Reddit posts; a Reflection Agent, which compares predicted outcomes against actual results; and finally, a Decision Maker, which issues trading signals based on the inputs from the other three agents [19]. They conducted back-testing on 2024 data and found that the strategy was only profitable during bull markets [19].

2.2 Deep Learning Strategies

Departing from traditional time-series models, Xiang et al. proposed a graph-based approach that leverages historical price data to construct dynamic relational graphs, employing a Graph Neural Network (GNN) for stock price prediction [20]. Their approach was back-tested on stocks comprising the S&P 500 and CSI 300 indices in 2020, adopting a daily buy-hold-sell strategy. Specifically, they employed the “top-k” selection method: buying stocks with the highest predicted returns, holding those with consistently strong predictions, and selling the remainder of the portfolio [20]. Their strategy achieved a 66.5% Annualised Return Rate (ARR) and a 1.421 Annualised Sharpe Ratio (ASR) in U.S. markets, and a 63.2% ARR with a 1.881 ASR in Chinese markets [20]. However, a limitation of GNN-based methods is that node representations may become overly similar, which can obscure hierarchical relationships, such as distinctions between individual and sector-level trends, potentially reducing predictive accuracy [21]. Additionally, live market testing appears to be limited to qualitative analysis, with no quantitative results reported.

In contrast, InvariantStock, proposed by Cao et al., tackles the challenge of unpredictable changes in market dynamics by learning invariant features through an environment-aware and environment-agnostic prediction model based on FactorVAE [22]. They construct an equally weighted, long-only portfolio of Chinese stocks using a top-*k* selection method [22]. When back-tested from January 2020 to October 2022, InvariantStock achieved an 83.15% ARR and a 3.7198 ASR [22]. The authors also tested the model on U.S. equities, permitting the short-selling of stocks with the lowest predicted returns [22]. However, performance declined in the U.S. back-test due to training on fewer features [22]. Additionally, the model’s reliance on historical data limits its adaptability in unprecedented market conditions, where non-historical factors may play a critical role [22]. Moreover, the absence of live market validation limits insights into the strategy’s practical performance.

2.3 Technical Analysis

Gerritsen et al. evaluated the use of technical trading rules in the Bitcoin market, including moving averages (MA), trading range breakouts, Moving Average Convergence Divergence (MACD), Rate of Change (ROC), On-balance Volume (OBV), Relative Strength Index (RSI), and Bollinger Bands (BB) to generate long or short trading signals [23]. Based on daily data from July 2010 to January 2019, they found that MA rules yielded similar returns to buy-and-hold, whereas the trading range breakout rule consistently outperformed it with a 1.78 ASR [23]. MACD and ROC rules only outperformed the benchmark when using the "Long or out of Bitcoin" strategy, whereas the RSI and BB rules consistently underperformed [23]. However, their study overlooks the impact of exchange fees and lacks live market testing [23].

In contrast, Gudapati et al. proposed two strategies using a combination of Exponential Moving Average (EMA), RSI, and Parabolic Stop and Reverse (PSAR) indicators [24]. They employed the use of EMA on RSI to classify the market as being in an uptrend, downtrend, or fluctuating state [24]. Following this, they strategically generate a trading signal selecting either EMA, RSI, or PSAR rules accordingly [24] (refer to A.2.2). Using daily Bitcoin prices from Yahoo Finance between 2018 and 2022 for back-testing, their most effective strategy yielded a total return of 701.77% [24]. However, their back-test overlooks exchange fees, lacks risk analysis, and their method omits live market validation.

Alternatively, Deprez et al. developed an extensive set of 18,840 technical trading rules for the BTC/USD market [25]. These rules span six widely used categories: MA, filter rules, support and resistance (S&R) rules, channel breakout (CB) rules, OBV, and RSI [25]. They evaluated these rules using data sampled at four different frequencies, ranging from daily to minute-level intervals [25]. Their results indicate that technical trading strategies can outperform a simple buy-and-hold approach, particularly in terms of risk-adjusted returns [25]. However, this advantage is less pronounced when looking at absolute returns, especially after factoring transaction costs [25]. Furthermore, incorporating margin trading and short selling can enhance returns but also elevate downside risk and trading costs [25].

2.4 Momentum Trading

Han et al. proposed a time-series and cross-sectional momentum trading strategy [26]. Under the time series method, a long position is taken if a portfolio's return (over a look-back period) falls in the top third of historical returns, a short position is taken if the return falls in the bottom third, with cash held otherwise [26]. Alternatively, under the cross-sectional method, coins are sorted into quintiles by return, with long positions taken in the top quintile and short positions in the bottom quintile simultaneously [26]. They back-tested their methods from 2013-2023 on hourly data, demonstrating profitability with high tail risk [26]. Short positions led to large losses, and long-only strategies were exposed to volatility, though their best strategy achieved a Sharpe ratio of 1.5 [26]. However, their results may be overly optimistic due to data-snooping bias, as multiple configurations were trialled on the same dataset, presenting a best-case scenario [26].

Alternatively, Chu et al. propose a signal-based high-frequency momentum trading strategy for Bitcoin, Ethereum, Dash, Litecoin, MaidSafeCoin, Monero, and Ripple (all USD-quoted), using EMA [27]. For the time-series method, buy or sell orders (dependant on signal polarity) are placed for an amount equal to the signal divided by 7 in USD worth of the cryptocurrency [27]. For the cross-sectional method, the top three signals are bought and the bottom three shorted, each with one-sixth of the signal value in USD [27]. They tested the strategy on both portfolios; however, a passive strategy outperformed their approach [27]. Their analysis is limited to back-testing a single bullish phase, and omits transaction costs [27].

2.5 Arbitrage

Gebbia et al. implemented a two-way arbitrage strategy for seven cryptocurrencies on Binance, Huobi, Lykke, Bitstamp and Kucoin [28]. They found that their method couldn't exploit arbitrage opportunities due to small spreads across the different exchanges, except for Bitstamp, where orders were quickly filled before the agent could place an order [28]. Furthermore, they employed both spatial and cyclical arbitrage across the DEXs Uniswap, Sushiswap, Pancakeswap, and Biswap [28]. Spatial arbitrage targets price discrepancies for the same asset across multiple DEXs, while cyclical arbitrage profits from mispricing within one DEX through sequential conversions [28]. They found that arbitrage opportunities were scarce, and that transaction fees on Ethereum blockchain-based DEXs were prohibitively high [28]. However, their scope was restricted to a small set of cryptocurrencies and exchanges, making a conclusive arbitrage assessment infeasible [28].

In contrast, Schwertfeger et al. proposed a triangular arbitrage trading algorithm between Centralised Exchanges (CEXs) and DEXs [29]. Two systems were deployed using DeFiChain (DFI) as an intermediary coin: the first, 'dMFP,' traded Solana (SOL) to capitalize on the DFI/USDT, DFI/SOL, and SOL/USDT cryptocurrency pairs, while the second, 'dD6K,' featured independent Ethereum and Solana arbitrage agents that shared liquidity in DFI and USDT holdings [29]. In back-testing, dMFP achieved a 24.91% net profit, while dD6K recorded a 33.35% net profit, from November 27 to December 31, 2023 [29]. However, their simulation neglected the effect of order placement on bid-ask spreads, competition from other traders and assumed sufficient liquidity to fill the orders [29]. Furthermore, their arbitrage strategy faces market risk from holding illiquid assets, which are crucial for enabling fast order execution across exchanges without rebalancing [29]. These cryptocurrencies can depreciate against fiat currency and cannot be hedged via shorting [29].

Alternatively, White proposed a sandwich arbitrage algorithm (also known as front-running), where a trader uses priority fees to strategically insert their own trades between others, manipulating DEX exchange rates to profit from the resulting price differences [30]. Simulating their method on the Ethereum blockchain across ETH/USDC, BTC/USDC, and ETH/BTC markets, they found that profit potential increases with the number of pending trades [30]. However, their method assumes that all blockchain trades occur on a single DEX governed by a constant product formula [30], where the product of the two asset quantities in the liquidity pool remains constant, determining the exchange rate based on the asset ratio [31]. Their method ignores transaction costs, the existence of limit orders, and assumes a fixed number of trades per block [30].

2.6 Pairs Trading

Fil et al. proposed two pairs trading algorithms, based on the cointegration and distance method [32]. They selected pairs exhibiting long-run correlations and then established long-run equilibrium relationships to restore balance between the pairs [32]. Their study utilized a dataset of 23 cryptocurrencies with hourly and daily price samples from 2018 [32]. The distance method, back-tested to hourly data, was the only strategy with positive returns, exhibiting a 2.87% monthly gain [32]. However, their back-testing overlooks short selling interest costs, is not optimized for margin cost efficiency, relies on a relatively short back-test period, and lacks live market validation [32, 33].

2.7 Derivatives Trading

Zou explored arbitrage opportunities between the Bitcoin spot and perpetual futures markets on Binance by leveraging the funding rates mechanism, where traders pay or receive the difference between the futures and spot price every 8 hours [34]. Arbitrage arises by buying Bitcoin on the spot market while simultaneously shorting an equivalent quantity of perpetual contracts [34]. Across eight 3-month back-tested periods from 2019 to 2021, returns ranged from 0.29% to 5.87% [34]. They identified arbitrage opportunities particularly when funding rates were positive and during market crashes [34]. However, their back-testing assumes zero transaction costs and lacks live market validation [34].

Alternatively, Chi and Hao proposed a volatility-spread trading strategy that takes long or short positions in Bitcoin and Ethereum call options based on the spread between daily Generalized Autoregressive Conditional Heteroskedasticity (GARCH) forecasted volatility and the options' implied volatility, while maintaining delta neutrality through hedging (arbitrage) [35]. The strategy uses daily spot price data from Binance and tick-level options data from Deribit, with spreads updated each tick [35]. Back-tests from September 2019 to March 2020 yielded returns of 28.5% for Bitcoin options and 52.2% for Ethereum options, net of position costs [35]. However, their method overlooks asymmetric volatility responses to past returns; later studies suggest Bitcoin and Ethereum now exhibit a leverage effect akin to equities, implying increased market efficiency and reduced strategy effectiveness [35-37]. Moreover, their strategy lacks live market validation.

Alternatively, Malik proposed a high-frequency trading strategy for Bitcoin perpetual futures, forecasting future contract prices using volatility and mean models [37]. The market direction is predicted from the difference between current and expected prices [37]. The conditional mean is estimated via a rolling window of returns, while conditional volatility is modelled using GARCH variants, Heterogeneous Autoregressive Model (HAR-RV), Random Forest, and Support Vector Machines (SVM) [37]. Back-tests on Binance data sampled at 5, 15, 30, and 60-minute intervals showed the 5-minute Exponential GARCH (EGARCH) model performed best [37]. Assuming 0.03% transaction fees, the strategy achieved a significant 151.08% ARR and 75.11 ASR from January 2020 to February 2023 [37]. However, live testing is required to validate these results.

3 METHODOLOGY

Our strategy combines a sentiment analysis indicator, derived from social media and news sources, with technical analysis indicators to improve entry and exit timing, as shown below:

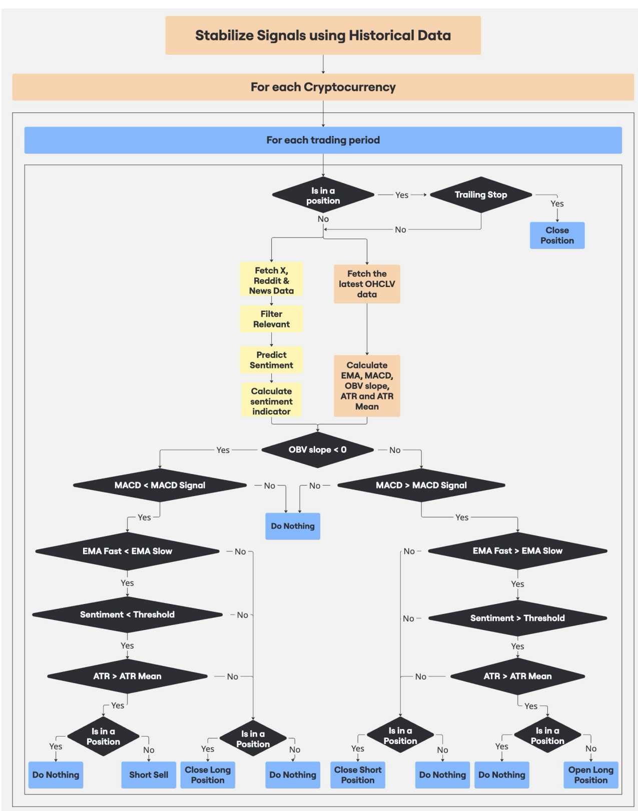

Figure 1 Flowchart of proposed strategy

Sentiment Analysis Indicator

Our sentiment analysis indicator incorporates posts from X, Reddit, and various news outlets. Text data is pre-processed by removing URLs, excess whitespace, hashtags, non-ASCII characters and "@" mentions. We then classify market relevance using DeepSeek-chat to reduce noise. The filtered data is once more passed through DeepSeek-chat to classify its sentiment (refer to A.3.1, A.3.2 and A.3.3).

We combine the sentiment of individual posts on X using the following formula, inspired by Rakotovanoelson [38]:

$$t_s = \sum_i^n p_i \times \frac{l_i+1}{f_i+1} \times (r_i + 1)$$

Where t_s represents the aggregated twitter sentiment, n denotes the number of posts (on the current day up to the current time), i denotes an individual post, p_i represents the individual sentiment prediction, l_i the number of likes, f_i the number of followers and r_i the number of re-tweets. For posts on reddit, we aggregate the sentiment using the formula below:

$$r_s = \sum_i^n p_i \times u_i$$

Where r_s represents aggregated reddit sentiment and u_i denotes the score of a post (net upvotes). We also combine news data using the formula below:

$$n_s = \sum_i^n p_i$$

Where n_s is the aggregated news sentiment. We aggregate and normalize sentiment across all three platforms according to the formula:

$$\text{sentiment} = \tanh(\,k \times (t_w \times t_s + r_w \times r_s + n_w \times n_s))$$

Where t_w, r_w and n_w are the weights of each source, and k is the scaling factor that adjusts the indicator's sensitivity to engagement levels. This yields a score between -1 (indicating negative sentiment) and 1 (indicating positive sentiment). Note, our sentiment indicator uses posts from the same date only, not a rolling 24-hour window.

Technical Analysis Indicators

In each trading period, we calculate EMA, MACD, OBV Slope, Average True Range (ATR), and the ATR mean for each cryptocurrency using TA-Lib [39] (see Appendix A.3.4 to A.3.8 for calculation details). These indicators are designed to capture trend, volatility, volume-based momentum, and trend-following momentum. To improve reliability, we include a time buffer to allow signals to stabilize before making decisions.

As illustrated in Figure 1, our strategy enters a long position when the OBV slope is greater than zero, indicating increasing buying pressure; the fast EMA crosses above the slow EMA, confirming a bullish trend; the MACD line crosses above the signal line, signalling bullish divergence; the sentiment indicator exceeds its threshold; and the ATR rises above its mean, suggesting a breakout from typical volatility levels. Conversely, the same is true for shorting, but the ATR must exceed the mean to confirm breakout. A position is closed when the OBV slope and MACD line cross the opposite thresholds, signalling a reversal in market pressure and momentum. For a long position, this occurs when the

OBV slope drops below zero and the MACD line crosses below the signal line; the opposite applies to close a short. The size of a position is determined by a fixed percentage of the fund's cash balance (in USDT).

Stop-Loss

A trailing-stop is applied to mitigate losses from adverse price movements. But more importantly, it also serves to lock in profits as the market moves in our favour. We allow only one open position per cryptocurrency, with a single trailing stop for each.

4 EXPERIMENTAL SETUP

4.1 Strategy Parameters

Traded Cryptocurrencies

We back-test our strategy on Dogecoin (DOGE), Cardano (ADA), Ripple (XRP), Ethereum (ETH), Solana (SOL), Bitcoin (BTC), and Monero (XMR), against USDT.

Back-testing periods

We define the in-sample back-testing period as the year 2022, and the out-of-sample back-testing period as the year 2024.

Position Size

For each trade signal, we allocate 20% of the fund's available cash balance (in USDT) to the position.

Stop-loss

We set a trailing-stop of 5%

MACD Indicator

The MACD timeframe parameters we consider for fast, slow, and signal lines are: (12, 26, 9), (8, 24, 9), and (5, 15, 9).

EMA

We consider a 20-period and 50-period slow EMA**,** with the closing price acting as the fast EMA.

ATR

We consider 10, 14, 20 and 50 period timeframes.

ATR Mean

We consider 14, 20 and 50 period timeframes

OBV Slope

We consider 10, 14, 20 and 50 period timeframes

Sentiment Threshold

A sentiment score above 0.7 is considered a bullish signal while a sentiment score below -0.5 is a bearish signal.

Sentiment Scale Factor

We use a scaling factor of $k = 0.000001$ for back-testing.

Sentiment Source Weighting

We combine various sentiment sources (X, Reddit and news), applying the following weighting scheme:

Source	t_w	r_w	n_w
Twitter + Reddit + News	0.4	0.4	0.2
News	0	0	1
Reddit	0	1	0
Twitter	1	0	0
News + Twitter	0.8	0	0.2
Reddit + Twitter	0.5	0.5	0

Fund Size

We back-test our strategies with an initial balance of 100,000 USDT.

Trading Period Length

We evaluate trading signals on 1-hour and 4-hour candlesticks.

4.2 Data Collection

Collection of X posts

We used the Apify web-scraping API to collect posts from X. For the in-sample period, tweets were scraped from the "Top" section, with filtering thresholds of 10 retweets, 20 likes, and 20 replies for DOGE, ADA, XRP, ETH, and BTC; 5 retweets, 10 likes, and 10 replies for SOL; and no filters for XMR. During the out-of-sample period, tweets were scraped from the "Latest" section, applying minimum like thresholds of 100 for ADA, 1000 for BTC, 20 for DOGE, 60 for ETH, and 5 for both SOL and XRP.

Collection of news headlines

We used the Crypto News+ dataset from Kaggle for in-sample back-testing [40].

Collection of reddit posts

We used the "Reddit - Crypto Posts" dataset from Kaggle for in-sample back-testing [41].

OHCLV Data

We used the Binance API to retrieve 1-hour and 4-hour K-line data for each cryptocurrency.

4.3 Evaluation Metrics

Win Rate

$$\text{Win Rate} = \frac{\text{Number of profitable trades}}{\text{Total nubmer of trades}}$$

The Win Rate (WR) represents the percentage of trades that were profitable; we use it to evaluate trading signal accuracy [42].

Annualised Sharpe Ratio

This metric represents the risk-adjusted return.

$$\text{Sharpe Ratio} = \frac{E(r_e - r_f)}{\sqrt{VAR(r_e - r_f)}} \times \sqrt{n}$$

Where r_e is the return each trading period, r_f is the risk-free return each trading period and n is the number of annual trading periods [43]. We use the average yield of a 3-month U.S. Treasury bill as a proxy for the risk-free rate [44]. For our in-sample testing period, we assume an annual risk-free rate of 2.09%, and for the out-of-sample period, 5.18%. [43].

Maximum Drawdown

$$\text{Maximum Drawdown} = \frac{\text{Trough Value} - \text{Peak Value}}{\text{Peak Value}}$$

Maximum Drawdown (MDD) is the peak-to-trough decline of an investment during a specific period [42]. We use MDD to evaluate the downside risk of different strategies [42].

Total Return

$$\text{Total Return} = \frac{\text{Final USDT balance}}{\text{Initial USDT balance}} - 1$$

Total Return (TR) is a percentage metric used to evaluate the profitability of a trading strategy.

5 RESULTS

5.1 In-sample Back-testing

Optimizing Parameters

Using in-sample data, we optimized our strategy parameters by running 48,384 combinations, evaluating each cryptocurrency independently. We then compiled the performance results from each back-test and calculated the following Pearson correlation matrix:

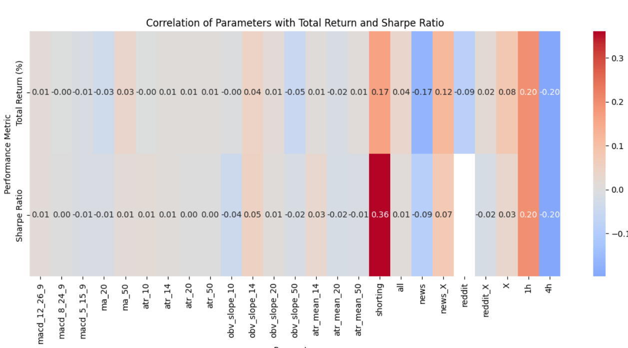

Figure 2 Pearson Correlation Matrix

As shown in Figure 2, strategies that incorporated short selling, hourly candlesticks, and leveraged either combined news and X sentiment or X-based sentiment alone exhibited the strongest positive correlation with higher returns and Sharpe ratios. A key limitation of the correlation matrix is that it evaluates relationships between individual parameters and performance in a pairwise manner, measuring the impact of changing one parameter while holding others constant. Consequently, it fails to capture complex interactions that emerge when multiple parameters vary simultaneously. This is particularly relevant for the parameters of the MACD, MA, ATR, ATR Mean, and OBV slope indicators, which may show weak individual correlations but could have a stronger combined effect when analysed together.

We selected the best configuration based on three objectives: maximizing the average Sharpe ratio, maximizing the average return, and minimizing the standard deviation of returns across all cryptocurrencies, thereby favouring a more generalized strategy. This selection process resulted in the following configuration:

MACD	Fast: 12, Slow: 26, Signal: 9
MA	Slow: 50, Fast: closing price
ATR	14
ATR Mean	50
OBV Slope	14
Allow Shorting	Yes
Sentiment Source	Twitter
Candlestick Timeframe	1 Hour

Performance Evaluation

We back-tested our best configuration using the in-sample portfolio achieving the following results:

Table 1 In-sample Performance of our strategy

Strategy	TR (%)	ASR	WR (%)	MDD (%)
Our Strategy	35.47	1.96	47.77	8.64
Buy-and-Hold	-11.73	-1.86	N/A	13.74

We found that our strategy performed exceptionally well, delivering high returns, a high Sharpe ratio, and a low maximum drawdown, indicating that the strategy is risk-efficient [45]. However, additional out-of-sample back-testing is necessary to determine whether the observed performance is robust or a result of overfitting. To better understand the drivers of in-sample performance, we further decomposed returns by cryptocurrency. The results, shown below, are based on back-tests conducted individually for each coin.

Table 2 Performance per Coin (Individual Back-tests)

Cryptocurrency	TR (%)	ASR	WR (%)	MDD (%)
ADA	2.70	0.15	41.67	4.6
BTC	1.69	-0.07	35.56	3.65
DOGE	**15.02**	**1.45**	54.84	3.44
ETH	3.32	0.36	57.89	**2.07**
SOL	8.53	1.14	**58.33**	2.95
XMR	0	0	0	0
XRP	1.73	-0.07	53.85	2.29

As expected, the meme coin DOGE generated the highest returns, while Bitcoin produced the lowest. This disparity likely arises from Dogecoin's speculative nature, fuelled by social media hype and limited fundamental support [46]. In contrast, Bitcoin is more mature and liquid, with increasing institutional participation, exhibiting time-varying efficiency in line with the Adaptive Market Hypothesis [47]. Periods of speculation and herding behaviour create opportunities for abnormal profits, though these are not constant, which may explain Bitcoin's comparatively lower returns [47]. We also note that our strategy did not execute any trades in the XMR/USDT market, resulting in a 0% return for that pair.

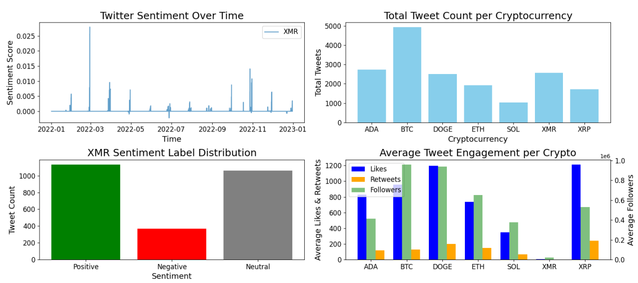

Figure 3 Composition of XMR sentiment signal

As shown in Figure 3, the sentiment signal for XMR remains below the trading threshold, resulting in no positions taken. This outcome is not due to tweet volume or sentiment distribution, as XMR has a substantial number of tweets and a positive sentiment bias. Rather, the low engagement metrics for XMR posts on X indicate limited market impact, justifying its exclusion from out-of-sample testing.

Correlation & Granger Causality

To investigate the impact of sentiment on price, we computed the Pearson and Spearman correlation coefficients, as well as Granger causality statistics (see A.5.1) across 24 lagged periods (i.e., 24 hours). Granger causality was deemed statistically significant when the p-value associated with the F-statistic was less than 5% [48]. Since these methods require stationary time series, we first transformed the raw price data into hourly returns [48]. We then applied the Augmented Dickey-Fuller (ADF) test to confirm the stationarity of both the return and sentiment data [48].

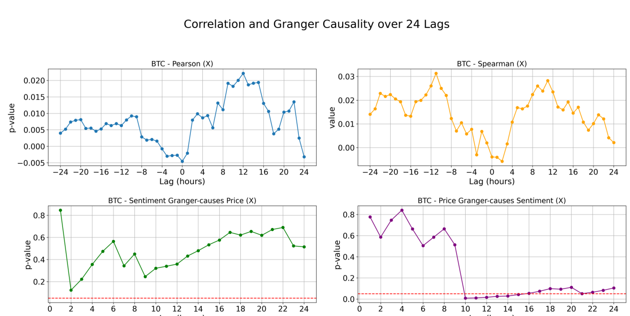

Figure 4 Correlation & Granger Causality of BTC

As shown in Figure 4, the Pearson and Spearman coefficients at positive lags reveal a very weak but mostly positive correlation between BTC returns and sentiment. Conversely, positive coefficients at negative lags indicate that changes in sentiment are weakly yet positively associated with changes in returns. Figure 4 shows that changes to sentiment do not Granger-cause returns, whereas changes in returns Granger-cause sentiment at lags between 10 and 16 periods (i.e., between 10 and 16 hours after the price change). Therefore, changes in Bitcoin sentiment observed on X cannot be used to improve return predictions (i.e. predict changes in price); however, they may serve as a lagging indicator to confirm existing trends.

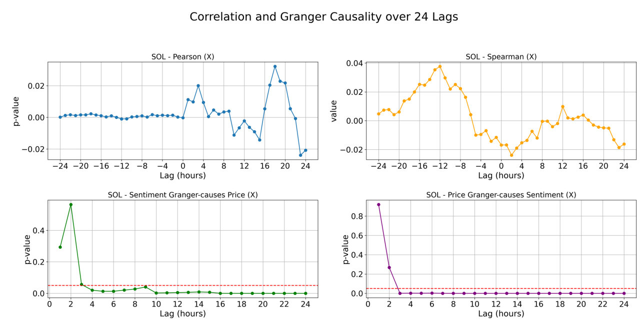

Figure 5 Correlation & Granger Causality of SOL

As illustrated in Figure 5 for SOL, the Pearson coefficient is weaker during the period when price leads sentiment, while the Spearman coefficient remains positive up to lag -5. Beyond this point, the Spearman coefficient becomes weakly negative, with a small positive deviation at lag 12, while the Pearson coefficient continues to fluctuate without clear trend. This disparity between the two measures suggests the presence of nonlinear relationships that Pearson, being limited to linear associations, may not capture. Overall, both coefficients are very weak for SOL, indicating a weak correlation. Analysis of Granger causality shows that beyond the third lag (3 hours), sentiment Granger-causes returns and returns Granger-cause sentiment. This suggests reflexive market behaviour in SOL/USDT, where sentiment and returns influence each other in a feedback loop [49, 50]. Such dynamics may indicate market inefficiency and present opportunities for profitable trading [49, 51].

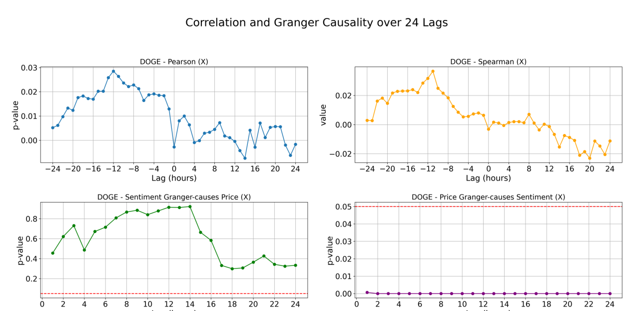

Figure 6 Correlation & Granger Causality of DOGE

Similarly, as illustrated in Figure 6, DOGE exhibits weak Pearson and Spearman correlations overall; notably, the Spearman correlation turns negative after the ninth lag, indicating a subtle but consistent inverse relationship applicable to older sentiment data. Importantly, sentiment does not Granger-cause returns at any lag, whereas returns Granger-cause sentiment across all tested lags, contrasting with the behaviour observed for BTC. The robustness of this directional causality is further supported by low p-values. Collectively, these findings reinforce the conclusion that sentiment primarily reflects past returns (price movements) rather than predicting future returns.
We note that the low correlation may be attributable to noise in hourly returns, and that correlation coefficients calculated from less frequently sampled data could exhibit greater significance. Furthermore, because our sentiment indicator resets daily, the two time-series may diverge at day boundaries, which could explain their weak correlation despite evidence supporting Granger causality.

5.2 Out-Of-Sample Back-testing

We compare our strategy against several benchmarks: Buy-and-Hold, a Cross-Sectional Momentum strategy, a Time-Series Momentum strategy, the Multi-Indicator Hierarchical Strategy (MIHS), the Multi-Indicator Hierarchical Constrained Strategy (MIHCS), and a baseline sentiment strategy employing market-driven labelling. The momentum strategies use value-weighted portfolios with a 1-day lookback and 7-day holding period for Cross-Sectional Momentum, and a 21-day lookback with a 7-day holding period for Time-Series Momentum, consistent with the observations of Han et al. [26]. The MIHS and MIHCS strategies apply a 7-day EMA on RSI to maximize performance, as described by Gudapati et al. [24]. For the benchmark sentiment strategy, we fine-tuned CryptoBERT on our in-sample data and optimized the market labelling parameters using a six-month rolling window preceding the out-of-sample period, to avoid look-ahead bias. The following table presents the results:

Table 3 Out-of-sample strategy returns

Strategy	TR (%)	ASR	WR (%)	MDD (%)
Our Strategy	11.04	0.57	41.51	**9.97**
Buy-and-Hold	5.91	0.12	N/A	11.9
Cross-Sectional Momentum	30.23	1.35	52.13	11.92
Time-Series Momentum	9.21	0.30	50.45	45.76
MIHS 7	**85.25**	**1.87**	48.84	19.08
MIHCS 7	74.04	1.76	39.13	16.94
Sentiment Baseline (Mean)	65.27	1.27	**58.19**	26.45
Sentiment Baseline (Majority)	50.95	1.00	54.07	19.72

As shown by the results in table 3, our strategy outperforms the time-series momentum strategy and buy-and-hold in terms of TR, ASR, and MDD. However, it underperforms other benchmarks in both TR and ASR. Notably, it falls well short of the 35.47% TR and 1.96 ASR achieved during in-sample back-testing, suggesting that overfitting or shifts in market dynamics may be limiting performance.

Market Dynamics

A potential shift in social sentiment, especially driven by changes in the popularity of hyped coins, could explain the lower performance. To better understand the drivers of sentiment, the following plot shows the average number of likes and retweets per cryptocurrency across both periods.

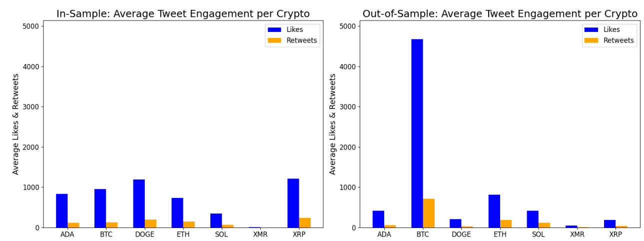

Figure 7 Tweet Engagement across both back-testing periods

Figure 7 illustrates that meme coin DOGE and altcoin XRP experienced a decline in engagement on X, whereas BTC saw a significant rise. Aside from shifts in sentiment dynamics, we also note that the in-sample period aligned with a bear market, whereas the out-of-sample period reflects a bull market.
To assess the robustness of the substantial returns generated by MIHS, MIHCS, and the Cross-Sectional Momentum strategy, we evaluated all three strategies during the in-sample bear market period. The results are presented in the table below.

Table 4 Bear Market Robustness Evaluation

Strategy	TR (%)	ASR	WR (%)	MDD (%)
MIHS 7	-33.30	-1.02	37.84	35.77
MIHCS 7	-29.03	-1.41	23.08	34.47
Cross-Sectional Momentum	**-10.79**	**-0.42**	**47.83**	**24.76**

The results suggest that these strategies lack robustness, having generated losses during the bear market period. Although MIHCS reports a smaller negative total return (and therefore greater excess return), it has a lower Sharpe ratio than MIHS due to a lower return volatility; that is, the returns are tightly clustered on the negative side. Nonetheless, the ASR for both strategies indicates that neither offers attractive risk-adjusted returns.

Performance Per Coin

To better understand the influence of market dynamics on performance, we back-tested our strategy on each coin individually.

Table 5 Out-of-sample performance per coin (Our Strategy)

Cryptocurrency	TR (%)	ASR	WR (%)	MDD (%)
ADA	**9.16**	**0.77**	44.44	1.98
BTC	-3.72	-1.69	37.11	7.91
DOGE	2.93	-0.41	45.45	3.87
ETH	1.89	-1.46	46.15	**1.47**
SOL	-1.56	-4.45	25.00	1.89
XRP	4.82	-0.03	**61.90**	2.29

An analysis of returns by cryptocurrency indicates that the altcoin ADA generated the highest return, while DOGE underperformed relative to its in-sample results. This divergence may be partially explained by a decline in social media engagement, as illustrated in Figure 7. However, although SOL also underperformed, its engagement metrics remained relatively stable, raising concerns about whether the conditions for reflexivity have been sustained. To assess market dynamics, Granger causality tests were conducted, and the results are presented below.

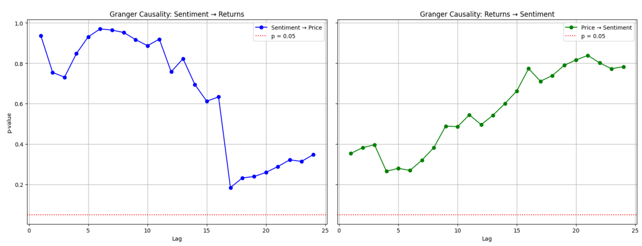

Figure 8 Granger Causality of SOL Returns and Sentiment

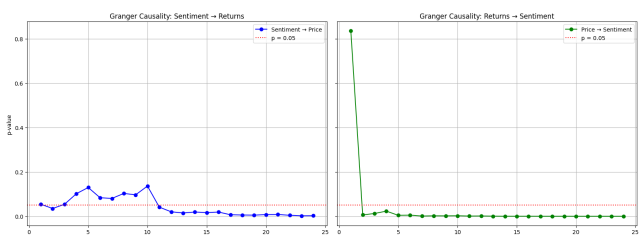

Figure 9 Granger Causality of ADA returns and Sentiment

As illustrated in Figure 8, sentiment does not Granger-cause returns, nor do returns Granger-cause sentiment. This absence of causality may help explain the observed underperformance of SOL during the out-of-sample period. Additionally, ADA exhibits market conditions similar to those of SOL during the in-sample period, which could account for its returns. However, other altcoins with comparable conditions still achieve positive, albeit underwhelming, performance (see Appendix 5.2), suggesting that other factors (such as technical indicator signals and engagement levels) may also influence returns. It is also possible that this observation is coincidental and unrelated.
To better understand the losses incurred by BTC, we plotted the signals generated by our strategy.

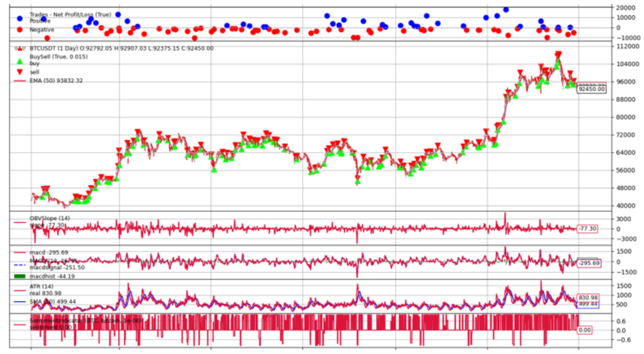

Figure 10 BTC Trading signals

As illustrated above, the rise in engagement for BTC has rendered the sentiment indicator overly sensitive, generating frequent (and possibly false) trading signals. This suggests that adjusting the engagement scaling factor k for BTC during the out-of-sample back-testing period may enhance strategy performance.

Evaluation of Technical Indicators (Ablation Study)

To evaluate the contribution of each technical indicator, we conduct an ablation study by removing one indicator at a time and assessing its impact. The results are presented below:

Table 6 Ablation study of strategy

Strategy	TR (%)	ASR	WR (%)	MDD (%)
Our Strategy	11.04	**0.57**	41.51	**9.97**
Without Sentiment	-61.46	-1.57	35.55	66.06
Without EMA	10.80	0.52	40.74	10.37
Without MACD	**11.10**	0.56	42.63	10.90
Without OBV Slope	5.32	0.06	**43.78**	13.82
Without ATR	-9.28	-1.27	37.84	20.18

As shown in Table 6, incorporating sentiment and ATR signals significantly improves performance, as their absence results in negative TRs and ASRs. Additionally, including the OBV slope nearly doubles the total return and substantially increases the ASR. In contrast, excluding EMA or MACD has a minimal effect, suggesting that the inclusion of both may be redundant.
To evaluate the predictive contribution of our sentiment indicator and assess whether the technical indicators offer maximal value, we conduct an additional ablation study using the benchmark method augmented with our sentiment indicator, as shown below:

Table 7 Sentiment indicator applied to benchmark strategy

Strategy	TR (%)	ASR	WR (%)	MDD (%)	Trades
Augmented Baseline	44.81	1.10	**58.33**	**15.16**	12
Baseline (Mean)	**65.27**	**1.27**	58.19	26.45	409
Baseline (Majority)	50.95	1.00	54.07	19.72	479

The augmented baseline outperforms our proposed strategy, suggesting that our choice of technical indicators is potentially constraining performance. Although the augmented baseline yields a slightly higher WR and lower MDD, it shows no improvement in TR or ASR compared to the mean baseline proposed by Moradi-Kamali et al. As our evaluation is limited to a bull market regime, further testing under bear and sideways conditions is needed to assess robustness. The low number of trades also indicates that our method employs stricter signal criteria, executing only 12 trades over the full back-test period. In contrast, the mean and majority-vote baselines each execute over 400.

5.4 Limitations

We employ Granger-causality to examine whether changes in sentiment can predict changes in returns. A key limitation of this approach is its assumption of a linear relationship, which may fail to capture the complex, nonlinear dynamics inherent in financial markets [52]. Additionally, the requirement for stationarity often necessitates converting price data into returns, which can obscure underlying trends [52].

Our approach assumes tweet sentiment reflects market direction, potentially oversimplifying the nuanced relationship between social media and price. Uncontextualized sentiment may lead to false signals, particularly in the presence of bias, misinformation, or sarcasm.

Another concern is the potential for look-ahead bias when using historical engagement metrics. Since posts often continue to attract likes and shares after a price movement event has occurred, the data may capture market reactions rather than preceding sentiment. This may lead to inflated return figures during back-testing.

Our back-testing framework also assumes no interest costs for short positions, which may result in overly optimistic performance metrics. Moreover, the out-of-sample testing was conducted during a predominantly bullish market regime. To assess robustness, further validation under bear and sideways markets is essential. Finally, we did not evaluate our strategy in live markets. Practical constraints such as lot sizes, margin interest, transaction fees, and engagement metrics (when used without look-ahead bias) may further diminish the performance observed in back-testing.

6 CONCLUSION

In conclusion, this study contributes to the field of algorithmic sentiment trading by (i) constructing a sentiment signal weighted by engagement metrics, (ii) implementing a sentiment analysis strategy integrated with technical indicators, (iii) exploring the application of large language models in trading algorithms, and (iv) hypothesizing a connection between the presence of Granger causality and market profitability.

Out-of-sample back-testing revealed that multi-indicator technical strategies (MIHS and MIHCS), while effective during bull markets, lacked robustness in bear market conditions. A similar pattern was observed for the cross-sectional momentum strategy. Our proposed strategy achieved a total return (TR) of 11.04% and an Annualised Sharpe ratio of 0.57, below the conventional threshold for risk-efficient strategies. In contrast, the benchmark sentiment strategy delivered significantly higher returns of 65.27% (mean-weighted) and 50.95% (majority-vote), with corresponding annual Sharpe ratios that met accepted standards for risk-adjusted performance [45]. However, the mean-weighted variant exhibited elevated maximum drawdowns [45]. Further analysis across different market regimes and live testing is required to conclusively validate its performance.

Our ablation study suggests that the performance of our sentiment strategy was constrained by the suboptimal selection of technical indicators. Notably, integrating our sentiment indicator into the benchmark strategy marginally improved the win rate and reduced the maximum drawdown.

6.1 Future Work

Our strategy predicts relevance and sentiment based on uncontextualized social media and news data. Future work could explore the incorporation of context-aware and temporal-context-aware data to improve prediction accuracy. Additionally, further research may compare the strategy's returns when leveraging other large language models, such as GPT or Gemini, as well as domain-specific models like FinGPT or FinLlama [53, 54]. Future research could also explore dynamically weighting sentiment sources to enhance the predictive power of the sentiment signal. Furthermore, our current approach focuses on sentiment from X, Reddit and news sources. Other platforms, including Telegram, Discord, TikTok, Instagram, and YouTube, offer promising opportunities to enrich sentiment signals and improve price predictability. Future research could also investigate whether incorporating sentiment signals enhances momentum-based trading strategies. Additionally, our use of a static trailing stop-loss presents an opportunity for refinement; dynamically adjusting it based on market volatility may improve risk management and returns. Finally, investigating adaptive position sizing techniques could enhance return potential while managing exposure to market risk.

7 LINK TO RELEASE

https://github.cs.adelaide.edu.au/a1887068/CryptoProject/releases/tag/Releasse

APPENDIX

A.1 Introduction

A.1.1 Factors influencing the price of a cryptocurrency

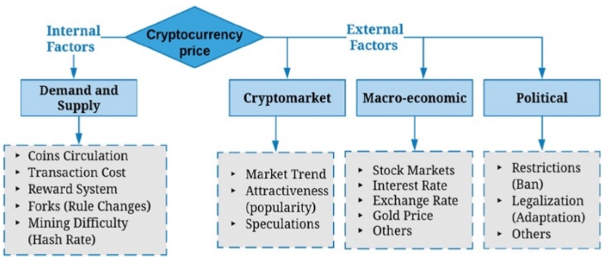

Figure A.1.1 Factors Affecting Cryptocurrency Prices [6]

A.2 Literature Review

A.2.1 Sentiment Analysis using "Market Driven Labelling"

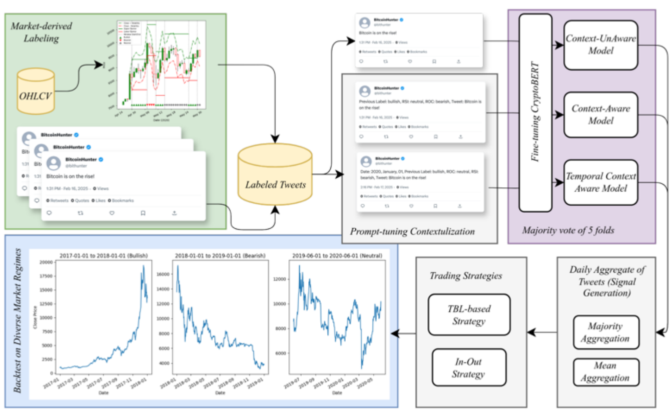

Figure A.2.1 Overall scheme of sentiment analysis method [18]

A.2.2 Multi-level technical indicator models

Gudapati et al. proposed two variations of a multi-indicator hierarchical strategy for technical analysis. The first, called the Multi-Indicator based Hierarchical Strategy (MIHS), is shown below**:**

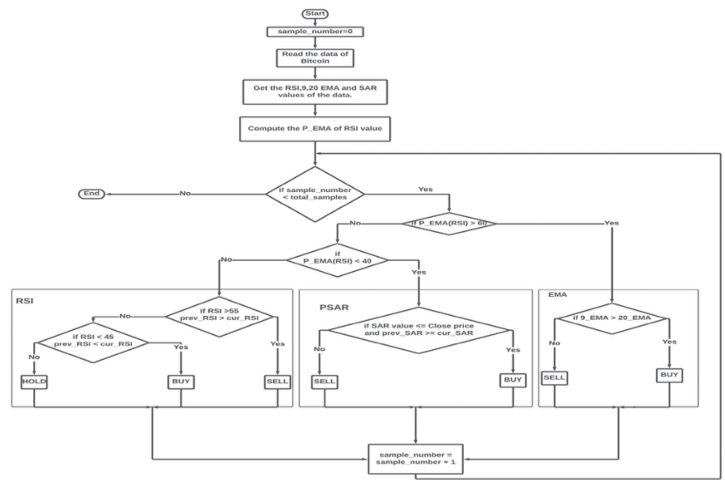

Figure A.2.2a Flowchart of MIHS logic [24]

The second variation, called the Multi-Indicator-based Hierarchical Constrained Strategy (MIHCS), is presented below:

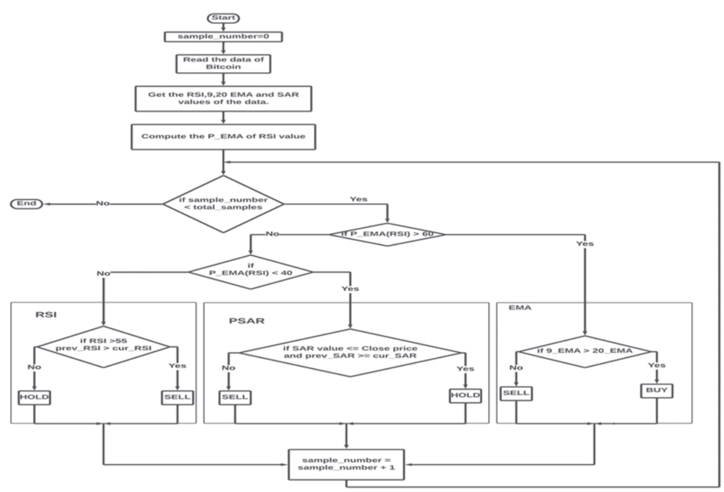

Figure A.2.2b Flowchart of MIHCS logic [24]

A.3 Methodology

A.3.1 DeepSeek Prompts for posts on X

We use the DeepSeek-Chat API to first classify the relevance of posts on X using the following configuration:

Role	Prompt
System	You are a financial relevance classifier. Classify tweets as market-relevant or not.
User	Determine if this tweet is related to financial markets, trading, investing, or cryptocurrency price movements. Text: "*X POST TEXT*" Respond only with "Yes" or "No".

Note that the system role defines the constraints and tone of the model's output, whereas the user role supplies the prompt to which the model responds. Following this, another API call is made to classify the sentiment using the following parameters:

Role	Prompt
System	Classify text as Positive, Negative, or Neutral.
User	Classify the sentiment of "*CRYPTOCURRENCY NAME*": "*X POST TEXT*" Respond only with one word: Positive, Negative, or Neutral.

This returns a Positive, Negative or Neutral label.

A.3.2 DeepSeek Prompts for Reddit Posts

Relevance Prompt:

Role	Prompt
System	You are a financial relevance classifier. Classify tweets as market-relevant or not.
User	Determine if this reddit post is related to financial markets, trading, investing, or cryptocurrency price movements. Text: "*REDDIT POST TEXT*" Respond only with "Yes" or "No".

Sentiment Analysis Prompt:

Role	Prompt
System	You are a sentiment analysis assistant. Classify text as Positive, Negative, or Neutral.
User	Classify the sentiment of "*TICKER*": "*REDDIT POST TEXT*" Respond only with one word: Positive, Negative, or Neutral.

A.3.3 DeepSeek Prompts for News Headlines

Sentiment Analysis Prompt:

Role	Prompt
System	You are a sentiment analysis assistant. Classify text as Positive, Negative, or Neutral.
User	Classify the sentiment of this text: "*NEWS HEADLINE*" Respond only with one word: Positive, Negative, or Neutral.

A.3.4 EMA Calculation

The EMA indicator typically uses the Simple Moving Average (SMA) for the first EMA period [23]. This is calculated using

$$MA_n = \frac{1}{n}\sum_{t=0}^{n} P_t$$

Where n is the length of the moving average period (same as length of EMA period in this case) and P_t is the closing price [23]. From this the EMA can be computed recursively using the following formula [23]:

$$EMA(P)_{t,n} = \left[P_t - EMA(P)_{t-1,n}\right] \times \frac{2}{n+1} + EMA(P)_{t-1,n}$$

Where t denotes the EMA period [23].

A.3.5 MACD

The MACD line is calculated according to the following [23]:

$$MACD(P)_t = EMA(P)_{t,s} - EMA(P)_{t,f}$$

Where s and f denote the length of the slow and fast EMA periods [23].

The MACD signal line is the EMA of the $MACD$ line:

$$SIGNAL(P)_t = \left[MACD(P)_t - EMA\big(MACD(B)\big)_{t-1,signal}\right] \times \frac{2}{\text{signal}+1} + EMA\big(MACD(P)\big)_{t-1,signal}$$

Where signal is the length of the signal period [23].

A.3.6 OBV Slope

To calculate the OBV slope, we must first calculate the OBV indicator. This is done as below:

$$OBV(P)_t = OBV(P)_{t-1} + \begin{cases} V \; if \; P_t > P_{t-1} \\ 0 \; if \; P_t = P_{t-1} \\ -V \; if \; P_t < P_{t-1} \end{cases}$$

Where V is the volume of transactions in the period [23]. We calculate the slope of the last n periods using the least squares method [39].

A.3.7 ATR

To calculate the ATR, the True Range (TR) must be defined first:

$$TR = \max[(H - L), |H - C_p|, |L - C_p|]$$

Where H is the high of the current period, L is the low of the current period and C_p is the closing price of the previous period [55]. The ATR is therefore an SMA of these measures

$$ATR = \frac{1}{n}\sum_{t}^{n} TR_t$$

Where n is the length of the ATR period, and t is the period [55].

A.3.8 Mean ATR

The Mean ATR is an SMA of the ATR

$$Mean\ ATR = \frac{1}{n}\sum_{t}^{n} ATR_t$$

Where n is the length of the Mean ATR period, and t is the period.

A.5 Results

A.5.1 Granger Causality of Coins during in-sample period

Figure 11 Granger Causality of the In-Sample Portfolio

A.5.2 Granger Causality of Coins during the out-of-sample period

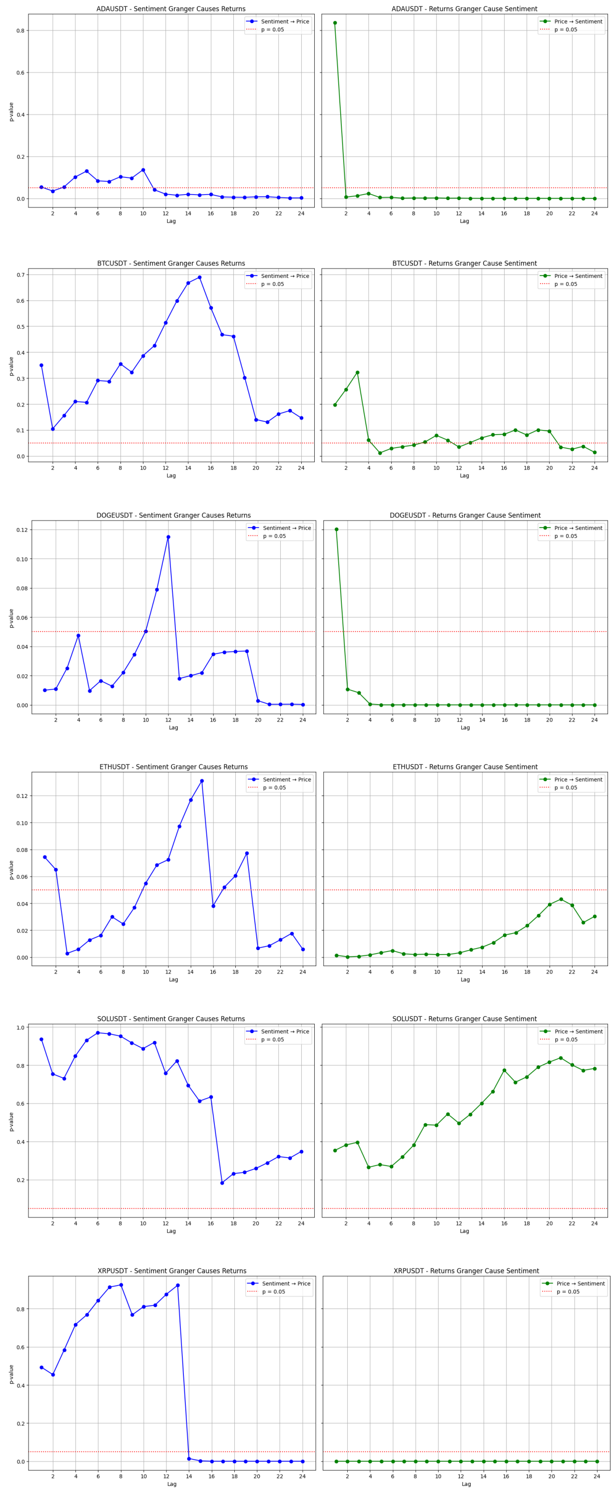

Figure 12 Granger Causality of the Out-of-Sample Portfolio

REFERENCES

[1] J. Sun, Y. Zhou, and J. Lin, "Using machine learning for cryptocurrency trading," in *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, 6-9 May 2019 2019, pp. 647-652, doi: 10.1109/ICPHYS.2019.8780358.

[2] C. DeVon. "This type of coin is 'among the riskiest of cryptocurrencies,' investing expert says—here's what to know." CNBC. (accessed Jun. 11, 2025).

[3] A. Hayes. "Technical Analysis: What It Is and How to Use It in Investing." Dotdash Meredith. https://www.investopedia.com/terms/t/technicalanalysis.asp (accessed Mar. 23, 2025).

[4] N. Reiff. "Cryptocurrency Exchanges: What They Are and How to Choose." Dotdash Meredith. https://www.investopedia.com/tech/190-cryptocurrency-exchanges-so-how-choose/ (accessed Mar. 25, 2025).

[5] "Why Fundamental and Technical Analysis Matter in Crypto." Nadcab Labs. (accessed Jun. 11 2025).

[6] A. M. Khedr, I. Arif, P. R. P V, M. El-Bannany, S. M. Alhashmi, and M. Sreedharan, "Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey," *Intelligent Systems in Accounting, Finance and Management,* vol. 28, no. 1, pp. 3-34, 2021, doi: https://doi.org/10.1002/isaf.1488.

[7] L. K. Felizardo *et al.*, "Outperforming algorithmic trading reinforcement learning systems: A supervised approach to the cryptocurrency market," *Expert Systems with Applications,* vol. 202, p. 117259, 2022, doi: https://doi.org/10.1016/j.eswa.2022.117259.

[8] A. Vo and C. Yost-Bremm, "A High-Frequency Algorithmic Trading Strategy for Cryptocurrency," *Journal of Computer Information Systems,* vol. 60, no. 6, pp. 555-568, 2020, doi: 10.1080/08874417.2018.1552090.

[9] J. Zhang, K. Cai, and J. Wen, "A survey of deep learning applications in cryptocurrency," *iScience,* vol. 27, no. 1, 2024, doi: 10.1016/j.isci.2023.108509.

[10] A. Shavandi and M. Khedmati, "A multi-agent deep reinforcement learning framework for algorithmic trading in financial markets," *Expert Systems with Applications,* vol. 208, p. 118124, 2022, doi: https://doi.org/10.1016/j.eswa.2022.118124.

[11] S. Otabek and J. Choi, "From Prediction to Profit: A Comprehensive Review of Cryptocurrency Trading Strategies and Price Forecasting Techniques," *IEEE Access,* vol. 12, pp. 87039-87064, 2024, doi: 10.1109/ACCESS.2024.3417449.

[12] M. Yousaf, M. Tariq, A. Jabbar, and S. Q. Jalil, "A Comprehensive Survey of Cryptocurrency Forecasting: Methods, Trends, and Challenges," in *Preprints*, ed: Preprints, 2024.

[13] S. Zhumagaziyev, "Can Reddit Sentiment Predict Bitcoin Returns?," Economics M.S. thesis, Department of Economics and Business, Central European University, Vienna, Austria, 2023. [Online]. Available: https://www.etd.ceu.edu/2023/zhumagaziyev_sh.pdf

[14] S. Agrawal, N. Kumar, G. Rathee, C. A. Kerrache, C. T. Calafate, and M. Bilal, "Improving stock market prediction accuracy using sentiment and technical analysis," *Electronic Commerce Research,* 2024/06/26 2024, doi: 10.1007/s10660-024-09874-x.

[15] L. Belcastro, D. Carbone, C. Cosentino, F. Marozzo, and P. Trunfio, "Enhancing Cryptocurrency Price Forecasting by Integrating Machine Learning with Social Media and Market Data," *Algorithms,* vol. 16, no. 12, p. 542, Nov. 2023, doi: 10.3390/a16120542.

[16] M. Kang, J. Hong, and S. Kim, "Harnessing technical indicators with deep learning based price forecasting for cryptocurrency trading," *Physica A: Statistical Mechanics and its Applications,* vol. 660, p. 130359, 2025, doi: https://doi.org/10.1016/j.physa.2025.130359.

[17] A. Sarkar, A. K. Sahoo, S. Sah, and C. Pradhan, "LSTMSA: A Novel Approach for Stock Market Prediction Using LSTM and Sentiment Analysis," in *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*, 13-14 March 2020, pp. 1-6, doi: 10.1109/ICCSEA49143.2020.9132928.

[18] H. Moradi-Kamali, M.-H. Rajabi-Ghozlou, M. Ghazavi, A. Soltani, A. Sattarzadeh, and R. Entezari-Maleki, "Market-Derived Financial Sentiment Analysis: Context-Aware Language Models for Crypto Forecasting," *ArXiv,* vol. abs/2502.14897, 2025.

[19] Q. Han, Q. Wang, A. Yoshikawa, and M. Yamamura, "PulseReddit: A Novel Reddit Dataset for Benchmarking MAS in High-Frequency Cryptocurrency Trading," 2025.

[20] S. Xiang, D. Cheng, C. Shang, Y. Zhang, and Y. Liang, "Temporal and Heterogeneous Graph Neural Network for Financial Time Series Prediction," presented at the Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, 2022. [Online]. Available: https://doi.org/10.1145/3511808.3557089.

[21] Z. You, Z. Shi, H. Bo, J. Cartlidge, L. Zhang, and Y. Ge, "DGDNN: Decoupled Graph Diffusion Neural Network for Stock Movement Prediction," *16th International Conference on Agents and Artificial Intelligence (ICAART),* vol. 2, pp. 431-442, 2024. [Online]. Available: https://arxiv.org/abs/2401.01846.

[22] H. Cao *et al.*, "InvariantStock: Learning Invariant Features for Mastering the Shifting Market," 2024. Accessed: Mar. 27, 2025. [Online]. Available: https://arxiv.org/abs/2409.00671

[23] D. F. Gerritsen, E. Bouri, E. Ramezanifar, and D. Roubaud, "The profitability of technical trading rules in the Bitcoin market," *Finance Research Letters,* vol. 34, p. 101263, 2020, doi: https://doi.org/10.1016/j.frl.2019.08.011.

[24] V. S. S. K. R. N. Gudapati, P. G, R. M, and A. V. Narasimhadhan, "Multi Indicator based Hierarchical Strategies for Technical Analysis of Crypto market Paradigm," *International Journal of Electrical and Computer Engineering Systems,* vol. 14, no. 7, pp. 765-780, 2023, doi: 10.32985/ijeces.14.7.4.

[25] N. Deprez and M. Frömmel, "Are simple technical trading rules profitable in bitcoin markets?," *International Review of Economics & Finance,* vol. 93, pp. 858-874, 2024, doi: https://doi.org/10.1016/j.iref.2024.05.003.

[26] C. Han, B. Kang, and J. Ryu, "Time-Series and Cross-Sectional Momentum in the Cryptocurrency Market: A Comprehensive Analysis under Realistic Assumptions," 2024. Accessed: Mar. 29, 2025. [Online]. Available: http://dx.doi.org/10.2139/ssrn.4675565

[27] J. Chu, S. Chan, and Y. Zhang, "High frequency momentum trading with cryptocurrencies," *Research in International Business and Finance,* vol. 52, p. 101176, 2020, doi: https://doi.org/10.1016/j.ribaf.2019.101176.

[28] C. Gebbia, "Analysis and Implementation of Arbitrage Bots in Centralized and Decentralized Finance," B.S. thesis, Department of Informatics, University of Zurich, Zurich, 2022. [Online]. Available: https://capuana.ifi.uzh.ch/publications/PDFs/22625_Analysis_and_Implementation_of_Arbitrage_Bots_in_Centralized_and_Decentralized_Finance.pdf

[29] L. Schwertfeger and B. Vogt, "Arbitrage trading between decentral and central cryptocurrency exchanges," 2024. Accessed: Mar. 27, 2025. [Online]. Available: http://dx.doi.org/10.2139/ssrn.5055292

[30] G. M. White, "Optimal Ordering to Maximize MEV Arbitrage," M.S., Brigham Young University, United States -- Utah, 30670346, 2023. [Online]. Available: https://www.proquest.com/dissertations-theses/optimal-ordering-maximize-mev-arbitrage/docview/2890690337/se-2?accountid=8203

[31] P. Tomar. "Constant Product Automated Market Maker: Everything you need to know." Medium. https://medium.com/@tomarpari90/constant-product-automated-market-maker-everything-you-need-to-know-5bfeb0251ef2 (accessed Mar. 26, 2025).

[32] M. Fil, "Pairs Trading in Cryptocurrency Markets," B.S. thesis, Faculty Of Social Sciences, Charles University, 2019. [Online]. Available: https://dspace.cuni.cz/bitstream/handle/20.500.11956/109634/130268190.pdf?sequence=1&isAllowed=y

[33] S. T. G. Nair, "Pairs trading in cryptocurrency market: A long-short story," *Investment Management and Financial Innovations,* vol. 18, no. 3, pp. 127-141, 2021, doi: 10.21511/imfi.18(3).2021.12.

[34] J. Zou, "Exploring Arbitrage Opportunities between the BTC Spot and Futures Market based on Funding Rates Mechanism," in *Proceedings of the 2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022)*, 2022: Atlantis Press, pp. 1923-1928, doi: 10.2991/aebmr.k.220307.316. [Online]. Available: https://doi.org/10.2991/aebmr.k.220307.316

[35] Y. Chi and W. Hao, "Volatility models for cryptocurrencies and applications in the options market," *Journal of International Financial Markets, Institutions and Money,* vol. 75, p. 101421, 2021, doi: https://doi.org/10.1016/j.intfin.2021.101421.

[36] Y. Chen, P. Bilokon, C. Hales, and L. Kerr, "Real-time VaR Calculations for Crypto Derivatives in KDB /Q," 2023. Accessed: Mar. 30, 2025. [Online]. Available: https://ssrn.com/abstract=4568236

[37] A. Malik, "A comparison of machine learning and econometric models for pricing perpetual Bitcoin futures and their application to algorithmic trading," *Expert Systems,* vol. 40, no. 10, p. e13414, 2023, doi: https://doi.org/10.1111/exsy.13414.

[38] A. Rakotovanoelson. "Decoding Twitter Sentiment and Cryptocurrency Price Movements." Medium. https://medium.com/@mister.anthony44/decoding-twitter-sentiment-and-cryptocurrency-price-movements-f9472e97a471 (accessed May. 10 2025).

[39] *TA-Lib: Technical Analysis Library*. (2025). [Online]. Available: https://ta-lib.org/

[40] Oliviervha. *Crypto News +*, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/oliviervha/crypto-news

[41] leukipp. *Reddit - Crypto Posts (r/cryptocurrency, r/eth...)*, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/leukipp/reddit-crypto-data/data

[42] Tradetron. "How to Evaluate the Performance of Algorithmic Trading Strategies." https://tradetron.tech/blog/how-to-evaluate-the-performance-of-algorithmic-trading-strategies (accessed Apr. 10 2025).

[43] QuantStart. "Annualised Rolling Sharpe Ratio in QSTrader." https://www.quantstart.com/articles/annualised-rolling-sharpe-ratio-in-qstrader/ (accessed Apr. 10 2025).

[44] FRED. "Market Yield on U.S. Treasury Securities at 3-Month Constant Maturity, Quoted on an Investment Basis." https://fred.stlouisfed.org/series/DGS3MO (accessed Apr. 20 2025).

[45] B. Young. "Top 5 Metrics for Evaluating Trading Strategies." LuxAlgo. https://www.luxalgo.com/blog/top-5-metrics-for-evaluating-trading-strategies/ (accessed Jun. 15 2025).

[46] S. Arun and G. Mandeep, "Exploring the potential of Dogecoin promoted by Elon Musk," *Scientific Journal of Metaverse and Blockchain Technologies,* vol. 2, no. 1, pp. 35-43, 02/29 2024, doi: 10.36676/sjmbt.v2.i1.06.

[47] K. H. Tan, "Is Bitcoin an Efficient Market? A Meta-Analytic Review of Price Dynamics and Efficient Market Hypothesis Tests," 2025.

[48] L. F. Pires, "Social Media Sentiment in Cryptocurrencies Markets: Application of granger causality and deep learning for price prediction," Finance M.S., Economics, University of Porto, 2021. [Online]. Available: http://sigarra.up.pt/fep/en/pub_geral.show_file?pi_doc_id=326632

[49] G. Soros, "Fallibility, reflexivity, and the human uncertainty principle," *Journal of Economic Methodology,* vol. 20, no. 4, pp. 309-329, 2013/12/01 2013, doi: 10.1080/1350178X.2013.859415.

[50] S. C. Long, Y. Xie, Z. Zhou, B. Lucey, and A. Urquhart, "From Whales to Waves: The Role of Social Media Sentiment in Shaping Cryptocurrency Markets," *Information Systems Research,* 2024, doi: http://dx.doi.org/10.2139/ssrn.4706410.

[51] J. Almeida and T. C. Gonçalves, "A systematic literature review of investor behavior in the cryptocurrency markets," *Journal of Behavioral and Experimental Finance,* vol. 37, p. 100785, 2023/03/01/ 2023, doi: https://doi.org/10.1016/j.jbef.2022.100785.

[52] A. Shojaie and E. B. Fox, "Granger Causality: A Review and Recent Advances," *Annual review of statistics and its application,* vol. 9 1, pp. 289-319, 2021.

[53] H. Yang, X.-Y. Liu, and C. Wang, "FinGPT: Open-Source Financial Large Language Models," *ArXiv,* vol. abs/2306.06031, 2023.

[54] T. Konstantinidis, G. Iacovides, M. Xu, T. Constantinides, and D. P. Mandic, "FinLlama: Financial Sentiment Classification for Algorithmic Trading Applications," *ArXiv,* vol. abs/2403.12285, 2024.

[55] A. Hayes. "Average True Range (ATR) Formula, What It Means, and How to Use It." Dotdash Meredith. https://www.investopedia.com/terms/a/atr.asp (accessed May. 5 2025).