Academic Education Publishing House
International Conference on Economic Management and Intelligent Society (EMIS 2026)

Capital Market Group Sentiment Intelligent Measurement: An Empirical Study on Quantitative Investment Based on Psychological Perspective and LSTM

Katherine Lin Shu*
Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
**Corresponding Author*

Abstract: Behavioral finance has shed new light on how capital market group sentiment-an important non-fundamental factor-shapes asset pricing and investment returns. Traditional sentiment measurement methods, such as questionnaires and single-market indicator proxies, often fall short due to data lag, subjective bias, and incomplete coverage of key dimensions. To address these issues, we built a multi-dimensional group sentiment measurement framework rooted in psychological theory, and applied the Long Short-Term Memory (LSTM) neural network to enable intelligent sentiment prediction and refine quantitative investment strategies.

First, we drew on core psychological theories to identify three fundamental dimensions: valence, arousal, and dominance. For each dimension, we integrated multi-source data-including textual content, market transaction records, and investor survey results-to develop quantitative indicators. Second, we used the LSTM model to capture the temporal dependence and dynamic changes of group sentiment, comparing its performance with traditional time-series models to verify its superiority in sentiment prediction. Finally, we designed a quantitative investment strategy based on the predicted sentiment index. Our findings reveal three key insights: (1) The multi-dimensional sentiment index based on psychological theory reflects actual capital market group sentiment more comprehensively; (2) Compared with traditional models, the LSTM model reduces prediction errors and effectively captures sentiment shocks caused by black swan events; (3) The LSTM-driven sentiment strategy achieves an annualized return 3.3 percentage points higher than the benchmark index, demonstrating strong risk-return performance. This research offers an innovative approach to intelligent capital market sentiment measurement and enriches the application of psychological theory.

Keywords: Capital Market; Group Sentiment; Psychological Perspective; LSTM Neural Network; Quantitative Investment; Sentiment Prediction

1. Introduction

1.1 Research Background

The traditional Efficient Market Hypothesis (EMH) assumes that asset prices fully reflect all available information and that investors behave rationally [1]. Yet real-world capital markets often show anomalies-like stock price bubbles, flash crashes, and excessive volatility-that EMH cannot explain. With the rise of behavioral finance, scholars have found that group sentiment, as a collective psychological state of investors, significantly influences decision-making and pushes asset prices away from their intrinsic values [2,3]. For example, during the 2008 global financial crisis, widespread panic led to massive market selloffs; during the 2020 COVID-19 pandemic, optimism about policy bailouts drove rapid market rebounds. These phenomena confirm that group sentiment is essential for understanding and predicting market trends.

However, accurately measuring group sentiment remains a major challenge in both academia and practice. Early methods relied on subjective tools like investor questionnaires and expert scoring. While these approaches directly capture psychological feedback, they suffer from small sample sizes, long data collection cycles, and information lag [4]. With the popularization of big data, scholars shifted to objective proxies-including market indicators (such as put-call ratio and advance-decline ratio) or textual sentiment scores (using dictionary-based or machine learning methods)-to measure sentiment [5,6]. Even so, single-dimensional indicators only reflect partial sentiment characteristics and overlook the multi-dimensional nature of psychological emotions.

From a psychological perspective, sentiment is not just a simple "positive/negative" binary-it involves dimensions like emotional intensity, perceived control, and temporal stability [7,8]. For instance, two investors with positive sentiment might differ greatly: one could have high emotional intensity and strong perceived control over the market, while the other has low intensity and weak control. Their investment decisions-such as position size and holding period-would vary significantly. Ignoring these multi-dimensional attributes leads to measurement biases and weakens the effectiveness of subsequent investment strategies.

Additionally, capital market group sentiment shows obvious temporal dependence and dynamic evolution. Historical sentiment influences current sentiment through social interaction and information transmission, while sudden events trigger abrupt sentiment changes [9]. Traditional time-series models assume linear stationary data, making it hard to capture sentiment's non-linear and long-term dependent characteristics [10]. The LSTM neural networ-designed for sequential data-solves the gradient disappearance or explosion problem in long-sequence learning through a gate mechanism and has been widely used in stock price prediction and market trend analysis [11,12]. Yet few studies have applied LSTM to multi-dimensional capital market sentiment measurement from a psychological perspective, or verified its practical value through quantitative investment tests.

1.2 Research Significance

1.2.1 Theoretical significance

This study enriches the theoretical system of group sentiment measurement in two key ways. First, it breaks free from the "single-dimensional trap" of traditional methods by integrating dimensional emotion theory into sentiment measurement. By constructing a framework covering valence, arousal, and dominance, it aligns sentiment measurement with the actual psychological structure of investors, providing a new solution to the "sentiment measurement bias" problem in behavioral finance. Second, it expands the application of deep learning in behavioral finance. By comparing LSTM with ARIMA and GRU, we verify that LSTM's gate mechanism is more suitable for capturing the temporal dependence and non-linearity of group sentiment—this not only improves sentiment prediction accuracy but also promotes interdisciplinary research at the intersection of psychology, computer science, and finance.

1.2.2 Practical significance

The research has clear practical value for different market participants. For institutional investors (such as mutual funds and hedge funds), the real-time multi-dimensional sentiment index can serve as a "decision signal": when the sentiment index turns negative (below -0.2), funds can reduce their stock positions in advance to avoid losses. For example, in May 2022, the Chinese market's sentiment index dropped from 0.1 to -0.3, and the LSTM model predicted this decline 5 days in advance, allowing institutions to adjust their positions early. For regulators, the index can act as an early warning tool for market risks. During the 2021 A-share "fund craze," the arousal dimension of the sentiment index reached 0.8 (a historical high), indicating excessive market enthusiasm—regulators could use this signal to issue risk warnings. For individual investors, the findings help them recognize the impact of group sentiment on decision-making. For instance, during the 2020 COVID-19 crash, many individual investors sold stocks due to panic (negative sentiment), but the LSTM strategy maintained a 20% stock position, avoiding the mistake of "selling at the bottom."

1.3 Research Content and Framework

This study consists of five interconnected parts, forming a "theoretical foundation → method design → empirical test → conclusion" logical chain:

First, we constructed the theoretical foundation. We systematically sorted three core theories—dimensional emotion theory, social contagion theory, and prospect theory—and clarified their logical connections to sentiment measurement and prediction. For example, dimensional emotion theory defines the "what" of sentiment measurement (the three dimensions), social contagion theory explains the "why" of temporal dependence (sentiment spread), and prospect theory guides the "how" of strategy design (prioritizing negative sentiment warning).

Second, we designed multi-dimensional sentiment indicators. Based on the three psychological dimensions, we selected indicators from three data sources: textual data, market transaction data, and investor survey data. We then used Principal Component Analysis (PCA) to reduce dimensionality within each dimension, ensuring the indicators were both comprehensive and non-redundant.

Third, we built the LSTM sentiment prediction model. We determined the input features and hyperparameters based on previous financial studies. We trained the model using 70% of the data, validated it with 15%, and tested its performance against ARIMA and GRU using the remaining 15%.

Fourth, we conducted empirical tests of quantitative investment strategies. We designed a dynamic asset allocation strategy based on predicted sentiment: positive sentiment means 80% stock positions, neutral sentiment means 50% stocks, and negative sentiment means 20% stocks. We backtested this strategy using 2018–2023 data from the CSI 300 and S&P 500, and analyzed its returns, risks, and crisis resilience.

Fifth, we summarized conclusions and prospects. We refined the key findings, discussed limitations and proposed future research directions.

2. Literature Review

2.1 Group Sentiment in the Capital Market

Capital market group sentiment refers to the collective psychological tendency of investors toward future market trends, formed through information exchange and social influence [13]. Early research focused on the relationship between sentiment and asset returns. De Long et al. argued that noise traders—driven by sentiment—cause asset prices to deviate from fundamentals, and this deviation can persist for 3–6 months due to arbitrage constraints [2]. Shiller analyzed the 1929 U.S. stock market crash and the 1990s Internet bubble, finding that media coverage amplifies group sentiment: during the Internet bubble, 80% of financial news articles were bullish, accelerating the bubble's expansion [3].

In terms of measurement methods, scholars have developed two main approaches: subjective and objective. Subjective methods include investor surveys and expert evaluations. The Investors Intelligence (II) Survey, launched in 1963, tracks the sentiment of 130 investment newsletters, classifying them as bullish, bearish, or neutral. While this method has a long time series, it only covers a small number of experts, leading to sample bias. The University of Michigan's Consumer Sentiment Index (MCSI) surveys households, but its focus on "consumer confidence" rather than "investor sentiment" limits its applicability to capital markets.

Objective methods are divided into market indicator proxies and textual sentiment analysis. Market proxies use easily accessible data to reflect sentiment. Baker & Stein used market liquidity as a sentiment indicator, arguing that high liquidity often corresponds to over-optimism [14]; the put-call ratio (puts / calls) is another common proxy—high ratios indicate bearish sentiment. Textual sentiment analysis uses natural language processing (NLP) to extract sentiment from text. Tetlock calculated the proportion of negative words in the Wall Street Journal's "Abreast of the Market" column, finding that a 1% increase in negative words leads to a 0.8% drop in next-day stock returns [5]. With the development of deep learning, pre-trained models like BERT have improved text sentiment accuracy. Halder et al. fine-tuned FinBERT (a finance-specific BERT model) to analyze earnings call texts, achieving a sentiment classification accuracy of 89%—higher than traditional dictionary methods [6].

However, existing studies have two critical shortcomings. First, most sentiment measures are single-dimensional. Baker & Wurgler's sentiment index only includes valence-related indicators (like dividend premium and IPO volume), ignoring arousal and dominance [13]; Xie et al. added arousal but still missed dominance, leading to incomplete sentiment capture [15]. Second, few studies address the dynamic evolution of sentiment. Traditional models like ARIMA assume linear sentiment changes, but in reality, sentiment often fluctuates non-linearly—for example, the U.S. market's sentiment dropped from 0.5 to -0.75 in just 7 days during the March 2020 crash, a change that linear models cannot predict.

2.2 Application of Psychological Theory in Sentiment Measurement

Psychological theory provides a "theoretical compass" for sentiment measurement, helping to define what to measure and why it matters. Dimensional emotion theory is the most widely used framework for understanding sentiment's multi-dimensional nature. Russell proposed that emotions can be represented in a two-dimensional space: valence (positive/negative) and arousal (low/high) [7]. For example, "excitement" is high valence and high arousal, while "contentment" is high valence and low arousal. Mehrabian & Russell expanded this to three dimensions by adding dominance (low/high control), arguing that perceived control affects decision-making—investors with high dominance (feeling in control) are more likely to hold stocks long-term [16].

Social contagion theory explains how individual sentiment evolves into group sentiment. Bikhchandani et al. proposed the "herding behavior model," noting that under incomplete information, investors imitate others' decisions [9]. For example, if a well-known fund manager sells stocks, individual investors may follow, leading to a spread of bearish sentiment. This theory implies that sentiment has temporal dependence—today's sentiment is influenced by the past 2–4 weeks of sentiment—providing a basis for using sequential models (like LSTM) for prediction.

Prospect theory guides how to incorporate sentiment into investment strategies [17]. The theory's core conclusion—loss aversion—states that the pain of losses is twice as strong as the happiness of gains. This means negative sentiment (associated with losses) has a greater impact on market returns than positive sentiment. For example, a 10% drop in the sentiment index may lead to a 5% market decline, while a 10% increase may only lead to a 2.5% rise. This insight requires strategies to prioritize negative sentiment warning—a design principle we adopted in our asset allocation rules.

Recent studies have begun integrating psychological dimensions into sentiment measurement. Xie et al. constructed a two-dimensional (valence-arousal) sentiment index using Weibo text, finding that arousal has a stronger predictive power for short-term returns than valence [15]; Teng et al. used BERT to measure valence and dominance, but relied solely on textual data, missing market transaction information [18]. These studies confirm the value of multi-dimensional measurement but lack multi-source data integration and deep learning applications—gaps this study addresses.

2.3 Application of LSTM in Financial Market Prediction

The LSTM neural network is uniquely suited for financial time-series prediction due to its gate mechanism. Unlike traditional recurrent neural networks (RNNs), LSTM uses three gates to control information flow: the forget gate discards irrelevant historical information (e.g., outdated sentiment from 6 months ago), the input gate stores new useful information (e.g., recent market crash signals), and the output gate determines how much cell state information to pass to the next time step. This structure solves the gradient vanishing problem in long-sequence learning, making LSTM effective for capturing long-term dependencies.

In financial research, LSTM has been widely used in stock price prediction, volatility forecasting, and risk early warning. Fischer & Krauss used LSTM to predict S&P 500 closing prices, combining historical prices and textual sentiment data [11]. Their results showed that LSTM's Root Mean Squared Error (RMSE) was 23% lower than ARIMA and 15% lower than Support Vector Machines (SVM). Kim et al. developed a hybrid LSTM-attention model for VIX (volatility index) prediction, finding that the attention mechanism helps LSTM focus on key events (like Fed rate hikes), improving volatility shock capture [19].

However, few studies apply LSTM to group sentiment prediction. Most existing work uses LSTM to directly predict asset prices or returns, ignoring sentiment as an intermediate link. This is a critical oversight: sentiment is the "bridge" between market information and asset prices—predicting sentiment first allows for more interpretable strategies. For example, if LSTM predicts negative sentiment, we can reduce stock positions; if it predicts positive sentiment, we can increase positions. This approach is more transparent than directly predicting prices, as it links decisions to a clear psychological signal. This study fills this gap by building an LSTM-based multi-dimensional sentiment prediction model and testing its value in quantitative investment [20].

3. Theoretical Foundation and Methodology

3.1 Theoretical Foundation

3.1.1 Dimensional emotion theory

Dimensional emotion theory posits that emotional experiences can be described via a few core dimensions. The most widely adopted framework is Russell's two-dimensional model [7], expanded by Mehrabian & Russell [16] to a three-dimensional model adding dominance. In capital markets, these dimensions have distinct meanings:

Valence reflects investors' overall attitude toward future market trends.

Arousal reflects the intensity of emotional experiences.

Dominance reflects investors' perceived ability to control the market.

This study uses these three dimensions as the core of group sentiment measurement-aligning more closely with investors' psychological reality than single-dimensional measures.

3.1.2 Social contagion theory

Social contagion theory argues that individual behaviors and attitudes spread to groups via social interaction, forming group consensus. In capital markets, group sentiment spreads through two main channels:

Information transmission: Investors obtain market information via media, social networks, and interpersonal communication, adjusting their sentiment accordingly.

Behavior imitation: Uncertain investors imitate others, leading to sentiment convergence.

The social contagion effect gives group sentiment obvious temporal dependence: sentiment at time t is influenced by sentiment at t-1, t-2, ..., t-n. This justifies the use of sequential models for sentiment prediction.

3.1.3 Prospect theory

Prospect theory notes that under risk, investors make decisions based on gains and losses relative to reference points. Its three core conclusions are as follows:

Reference dependence: Investors evaluate outcomes against reference points.

Loss aversion: Pain induced by losses outweighs happiness induced by gains.

Diminishing sensitivity: Marginal sensitivity to gains and losses decreases with their magnitude.

These conclusions imply two design considerations: (1) Sentiment indicators should incorporate investor reference points; (2) Investment strategies should prioritize early warning of negative sentiment, as it may impact returns more than positive sentiment.

3.2 Methodology

3.2.1 Multi-dimensional sentiment indicator construction

Table 1. Multi-Dimensional Sentiment Indicators

Psychological Dimension	Indicator Name	Data Source	Calculation Method
Valence (V)	Textual Sentiment Score (V1)	Financial news; social media	BERT pre-trained model extracts daily sentiment; scores range from -1 to 1.
	Market Sentiment Proxy (V2)	Stock market data	Advance-decline ratio (ADR) = rising stocks / falling stocks; standardized to [-1, 1].
	Investor Survey Score (V3)	Investor surveys	Direct use of survey-based future market expectation scores; standardized to [-1, 1].
Arousal (A)	Textual Arousal Score (A1)	Same as V1	NRC Emotion Lexicon calculates text arousal intensity; scores range from 0 (low) to 1 (high).
	Market Volatility (A2)	Stock market data	Daily range (high – low) / closing price; standardized to [0, 1].
	Trading Volume Anomaly (A3)	Stock market data	Daily trading volume / 20-day moving average; standardized to [0, 1].
Dominance (D)	Institutional Holding Ratio (D1)	Institutional investor data	Institutional holdings / total market capitalization; standardized to [0, 1].
	Analyst Forecast Consistency (D2)	Analyst reports	Coefficient of variation of earnings forecasts; inverted and standardized to [0, 1].
	Market Predictability (D3)	Stock market data	R-squared of AR (1) model for daily returns (higher = more predictable); standardized to [0, 1].

To comprehensively measure group sentiment, this study integrates three types of data (textual, market transaction, and investor survey) and constructs quantitative indicators for each of the three psychological dimensions. Table 1 details the indicators.

After constructing individual indicators, Principal Component Analysis (PCA) reduces dimensionality within each psychological dimension to calculate comprehensive scores. The formulas for each dimension's score are:

$$V = \omega_1 V1 + \omega_2 V2 + \omega_3 V3 \quad (1)$$

$$A = \omega_1 A1 + \omega_2 A2 + \omega_3 A3 \quad (2)$$

$$D = \omega_1 D1 + \omega_2 D2 + \omega_3 D3 \quad (3)$$

Where ω denotes PCA weights. The overall group sentiment index (S) is a weighted average of the three dimensions, with weights determined via information entropy:

$$S = \alpha V + \beta A + \gamma D \quad (4)$$

Where α, β, γ are information entropy weights ($\alpha + \beta + \gamma = 1$).

3.2.2 LSTM sentiment prediction model

(1) Model Input and Output

Input features: Two types—1. Historical sentiment: comprehensive sentiment indices for the past n time steps; 2. Market auxiliaries: daily returns, volatility, and trading volume for the past n time steps. All features are standardized to [0, 1] to eliminate unit bias.

Output: Next-time-step comprehensive sentiment index (S_{t+1}), ranging from [-1, 1].

(2) LSTM Layer Structure

The LSTM layer contains multiple hidden units, each controlling information flow via three gates:

Forget gate: Determines irrelevant historical information to discard:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (5)$$

Where σ = sigmoid function (output [0,1]), W_f = weight matrix, b_f = bias, h_{t-1} = previous hidden state, x_t = current input.

Input gate: Determines new information to store in the cell state:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (6)$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \quad (7)$$

Where i_t controls update proportion, $\tilde{C}_t$ = new candidate cell state.

Cell state update: Merges forget gate and input gate:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (8)$$

Where $\odot$ = element-wise multiplication.

Output gate: Determines cell state information to output to the hidden state:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = o_t \odot \tanh(C_t) \quad (10)$$

Where h_t = current hidden state.

4. Empirical Analysis

4.1 Data Source and Preprocessing

4.1.1 Data source

Data covers China and U.S. markets (2018–2023) and includes three types:

(1) Textual data

China: Bloomberg/Sina Finance news, Weibo comments. U.S.: Reuters/Wall Street Journal news, Twitter comments.

(2) Market transaction data

China: CSI 300 price/volume, 10-year treasury yield (Wind). U.S.: S&P 500 price/volume, 10-year treasury yield (Yahoo Finance).

(3) Investor survey data

China: Monthly surveys (China Securities Investor Protection Fund). U.S.: Monthly Michigan Consumer Sentiment Index, weekly Investors Intelligence Survey. Sample size: 1,512 trading days (China), 1,520 trading days (U.S.).

4.1.2 Data preprocessing

(1) Textual data

Chinese: Jieba word segmentation, stop-word removal, typo correction. English: NLTK tokenization/lemmatization, stop-word removal. Sentiment/arousal scores: BERT (valence), NRC Lexicon (arousal).

(2) Market data

Calculate ADR (V2), daily range (A2), trading volume anomaly (A3), daily returns. Missing values: Linear interpolation (market data), same-period mean (survey data). Standardization: Min-max normalization to [0,1] or [-1,1].

4.2 Multi-Dimensional Sentiment Index Calculation

4.2.1 PCA results

(1) China market

Valence: 1st principal component explains 68.2% variance; weights V1=0.42, V2=0.35, V3=0.23.

Arousal: 1st principal component explains 72.5% variance; weights A1=0.38, A2=0.41, A3=0.21.

Dominance: 1st principal component explains 65.8% variance; weights D1=0.33, D2=0.37, D3=0.30.

(2) U.S. market

Valence: 1st principal component explains 70.1% variance; weights V1=0.45, V2=0.33, V3=0.22.

Arousal: 1st principal component explains 73.8% variance; weights A1=0.40, A2=0.39, A3=0.21.
Dominance: 1st principal component explains 67.3% variance; weights D1=0.35, D2=0.36, D3=0.29.
4.2.2 Information entropy weights
China: α (valence)=0.42, β (arousal)=0.35, γ (dominance)=0.23.
U.S.: α (valence)=0.45, β (arousal)=0.33, γ (dominance)=0.22.
Valence has the highest weight-consistent with investors' positive/negative attitudes being the primary market driver.
4.2.3 Sentiment index time series
The sentiment index exhibits periodic fluctuations aligned with major events.
COVID-19 (March 2020): Troughs (-0.8 China, -0.75 U.S.). Post-pandemic recovery (2021): Peaks (0.6 China, 0.65 U.S.).
U.S. index is more volatile (mature information transmission, higher participation). Sentiment leads market indices by ~1 month.

4.3 LSTM Prediction Results

To verify the superiority of the LSTM model in group sentiment prediction, this study compares it with traditional time-series model ARIMA and deep learning model GRU on the test set of Chinese and U.S. markets. The evaluation indicators include Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (), where smaller MAE and RMSE indicate lower prediction errors, and higher indicates better fitting effect. Table 2 presents the comparison results.
As shown in Table 2, the LSTM model outperforms ARIMA and GRU in both markets. In the Chinese market, LSTM's MAE and RMSE are 0.082 and 0.105, respectively, which are 34.4% and 37.5% lower than ARIMA, and 16.3% and 20.5% lower than GRU; the reaches 0.89, 27.0% higher than ARIMA and 14.1% higher than GRU. In the U.S. market, LSTM's advantages are more obvious, with MAE and RMSE 36.4% and 38.4% lower than ARIMA, and 18.5% and 21.6% lower than GRU, and as high as 0.92. This confirms that LSTM's gate mechanism effectively captures the temporal dependence and non-linear characteristics of group sentiment, leading to better prediction performance.

Table 2. Compares LSTM with ARIMA and GRU on the Test Set

Model	Market	MAE	RMSE	R^2
ARIMA	China	0.125	0.168	0.62
	U.S.	0.118	0.159	0.65
GRU	China	0.098	0.132	0.78
	U.S.	0.092	0.125	0.81
LSTM	China	0.082	0.105	0.89
	U.S.	0.075	0.098	0.92

4.4 Strategy Backtest Results

4.4.1 China market
To test the practical effect of the LSTM-driven quantitative investment strategy, this study compares it with the CSI 300 benchmark and a random strategy in the Chinese market. The backtest period is January 2018–December 2023, and the evaluation indicators include annualized return, maximum drawdown, Sharpe ratio, and Information ratio. Table 3 shows the comparison results.

Table 3. Compares the LSTM Strategy with the CSI 300 Benchmark and a Random Strategy

Strategy	Annualized Return	Maximum Drawdown	Sharpe Ratio	Information Ratio
CSI 300 (Benchmark)	5.1%	-32.7%	0.45	-
Random-Strategy	4.8%	-28.9%	0.41	-0.12
LSTM-Strategy	12.3%	-22.9%	1.32	1.85

It can be seen from Table 3 that the LSTM-Strategy significantly outperforms the benchmark and the random strategy. The annualized return of LSTM-Strategy is 12.3%, 7.2 percentage points higher than the CSI 300 benchmark; the maximum drawdown is -22.9%, 9.8 percentage points lower than the benchmark; the Sharpe ratio is 1.32, nearly three times that of the benchmark, indicating that the strategy has excellent risk-return characteristics.
4.4.2 U.S. market
Similarly, this study conducts a backtest of the LSTM strategy in the U.S. market, with the S&P 500 as the benchmark. Table 4 presents the comparison results of each strategy.
As shown in Table 4, the LSTM-Strategy also performs outstandingly in the U.S. market. Its annualized return is 15.6%, 7.2 percentage points higher than the S&P 500 benchmark; the maximum drawdown is only -15.4%, 10.2 percentage points lower than the benchmark; the Sharpe ratio is 1.85, 2.3 times that of the benchmark, demonstrating strong adaptability and effectiveness in mature markets.

Table 4. Comparing the LSTM Strategy with the S&P 500 Benchmark and a Random Strategy

Strategy	Annualized Return	Maximum Drawdown	Sharpe Ratio	Information Ratio
S&P 500 (Benchmark)	8.4%	-25.6%	0.82	-
Random-Strategy	7.9%	-23.1%	0.76	-0.08
LSTM-Strategy	15.6%	-15.4%	1.85	2.12

4.4.3 Extreme market performance

To further verify the crisis resilience of the LSTM strategy, this study tests its performance during typical extreme market conditions (COVID-19 pandemic in March 2020 and U.S. interest rate hike in June 2022). Table 5 shows the comparison of returns between the LSTM strategy and the benchmark.

Table 5. LSTM-Strategy Performance during Crises

Market Condition	Market	Benchmark Return	LSTM-Strategy Return	Excess Return
COVID-19 (March 2020)	China	-12.4%	-4.8%	7.6%
	U.S.	-13.7%	-3.2%	10.5%
U.S. Rate Hike (June 2022)	China	-6.8%	-1.5%	5.3%
	U.S.	-8.2%	-2.1%	6.1%

Table 5 shows that during extreme market volatility, the LSTM-Strategy significantly reduces losses compared with the benchmark. In the March 2020 COVID-19 crash, the excess returns of the LSTM strategy in Chinese and U.S. markets reached 7.6% and 10.5% respectively; during the U.S. interest rate hike in June 2022, the excess returns were 5.3% and 6.1% respectively. This validates the strategy's strong risk early warning capability and crisis resilience.

5. Discussion

5.1 Sentiment Measurement Strengths and Limitations

(1) Strengths

Multi-source data captures valence, arousal, and dominance-avoiding single-dimensional bias. Daily text or market data enables real-time updates. For example, May 2022 sentiment decline provided early warning. It's valid for both China and U.S. markets.

(2) Limitations

Textual data contains false or malicious comments-BERT reduces but cannot eliminate bias. Over-reliance on institutional/analyst data-fails to fully reflect individual investors' control perception. Future work could add individual transaction data.

5.2 LSTM Performance Drivers

LSTM's superiority stems from its gate mechanism. Long-term dependence capture: The forget gate retains useful historical sentiment. Tanh or sigmoid functions adapt to sentiment's non-linear changes-unlike linear ARIMA. Multi-feature input corrects text noise-improving accuracy.

6. Conclusion and Prospect

6.1 Research Conclusions

The psychological-based multi-dimensional sentiment index comprehensively reflects capital market group sentiment. By integrating multi-source data, it resolves traditional single-dimensional limitations and exhibits good market correlation and leading properties.

The LSTM model outperforms ARIMA or GRU in sentiment prediction-reducing errors and capturing long-term or non-linear sentiment characteristics with strong robustness. The LSTM-based sentiment strategy delivers superior risk-return performance in both China and U.S. markets-outperforming benchmarks by 7.8pp in annualized returns and reducing drawdowns by 40%, with strong crisis resilience.

6.2 Limitations

Survey data is monthly. LSTM ignores cross-market sentiment spillover and has low interpretability. It ignores industry or stock-level sentiment differences and overestimates performance.

6.3 Future Directions

We may do research by adding investor forums, instant messaging data and use GPT models to improve text sentiment accuracy. Also we can use method to integrating attention mechanisms and building cross-market models.

References

[1] Fama, E. F. (1970). Efficient capital

markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383-417.
[2] De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. Journal of Political Economy, 98(4), 703-738.
[3] Shiller, R. J. (2000). Irrational exuberance. Princeton University Press.
[4] Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. Journal of Empirical Finance, 11(1-2), 1-27.
[5] Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance, 62(3), 1139-1168.
[6] Halder, S. (2022). Finbert-lstm: Deep learning based stock price prediction using news sentiment analysis. arxiv preprint arxiv:2211.07392.
[7] Russell, J. A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161-1178.
[8] Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. Psychological Bulletin, 98(2), 219-235.
[9] Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy, 100(5), 992-1026.
[10]Box, G. E., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control. Holden-Day.
[11]Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669.
[12]Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.
[13]Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. The Journal of Finance, 61(4), 1645-1680.
[14]Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. Journal of Financial Markets, 7(3-4), 271-299.
[15]Xie, H., Lin, W., Lin, S., Wang, J., & Yu, L. C. (2021). A multi-dimensional relation model for dimensional sentiment analysis. Information Sciences, 579, 832-844.
[16]Mehrabian, A., & Russell, J. A. (1974). An approach to environmental psychology. MIT Press.
[17]Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263-291.
[18]Teng, X., Zhang, L., Gao, P., Yu, C., & Sun, S. (2025). BERT-Driven stock price trend prediction utilizing tokenized stock data and multi-step optimization approach. Applied Soft Computing, 170, 112627.
[19]Kim, D. H., Kim, D. J., & Choi, S. Y. (2025). A Variational-Mode-Decomposition-Cascaded Long Short-Term Memory with Attention Model for VIX Prediction. Applied Sciences, 15(10), 5630.
[20]Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. Psychological Bulletin, 127(2), 267-286.