


Article

A Rule-Based Stock Trading Recommendation System Using Sentiment Analysis and Technical Indicators

Yuri Kim ¹, Sujin Yoo ² and Seongbin Park ^{1,*}

¹ Department of Computer Science, Korea University, 145, Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea; fhrml1004@korea.ac.kr

² Division of Computere Engineering, College of IT Engineering, Hansung University, Samseongyo-ro 16 gil, Seongbuk-gu, Seoul 02876, Republic of Korea; sujin.yoo@hansung.ac.kr

* Correspondence: hyperspace@korea.ac.kr

Abstract: This paper presents a stock trading recommendation system that integrates news sentiment analysis with the relative strength index (RSI) to provide informed buy–sell decisions. The system uses a rule-based natural language processing (NLP) approach to analyze recent news articles and combines the resulting sentiment scores with the RSI, which tracks stock momentum. By evaluating seven days of news data, the system assigns a sentiment score (1 to 100) that reflects market sentiment, while the RSI identifies overbought or oversold conditions. This combined approach allows traders to make data-driven buy, sell, or hold decisions in real time. In this study, we conducted a comparative study with benchmark indices across various subsets of stocks to evaluate their relative performance, highlighting our system’s competitive edge in terms of accuracy, profitability, and lightweight design with low computational cost. The results showed the system’s adaptability across different market segments and its potential to enhance trading outcomes. By integrating real-time sentiment analysis with technical indicators, the system offers a practical and actionable investment strategy.

Keywords: recommendation system; sentiment analysis; natural language processing



Academic Editor: Arkaitz Zubiaga

Received: 10 January 2025

Revised: 7 February 2025

Accepted: 12 February 2025

Published: 17 February 2025

Citation: Kim, Y.; Yoo, S.; Park, S. A Rule-Based Stock Trading Recommendation System Using Sentiment Analysis and Technical Indicators. *Electronics* **2025**, *14*, 773. <https://doi.org/10.3390/electronics14040773>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ultimate goal of stock trading is to generate profit by buying and selling company shares, with traders aiming to capitalize on price fluctuations that occur in the market [1]. However, stock prices are influenced by numerous factors, creating a complex system that makes it difficult to consistently generate profit [2]. The stock market’s highly volatile and interconnected nature adds to the challenge, making traditional time series modeling approaches less effective [3]. As a result, various strategies and tools have been developed over time, with quantitative methods playing a significant role in the evolution of capital markets.

In recent years, natural language processing (NLP) has gained increasing importance in quantitative financial strategies for three primary reasons: handling large volumes of unstructured data, predicting market sentiment through sentiment analysis, and enhancing automated trading systems [4–6]. The financial sector produces vast amounts of unstructured data, such as news articles, social media posts, earnings reports, and regulatory filings. NLP is a valuable tool for extracting meaningful insights from these data, supporting better decision-making [7]. Sentiment analysis, an NLP technique, helps traders gauge market sentiment, investor behavior, and public opinion by analyzing text from various sources. This analysis allows traders to evaluate how positive or negative sentiments

could influence stock prices or market trends. Additionally, NLP facilitates algorithmic trading by processing real-time financial news and adjusting trading strategies based on those findings.

However, while NLP-based stock trading systems offer significant advantages, they also face challenges. These include over-reliance on historical and sentiment data, high computational costs, and limited accessibility for individual investors, due to expensive hardware requirements [8]. Existing systems often require high initial investments and may not be suitable for personal use, due to substantial transaction fees and misleading claims about performance [9].

The research problem addressed in this study is as follows: How can a lightweight and computationally efficient stock trading recommendation system be developed to combine real-time news sentiment analysis with technical indicators, providing accurate and accessible buy–sell decisions for individual investors?

To address these challenges, this paper proposes a stock trading system that integrates NLP-based sentiment analysis of financial news with the relative strength index (RSI), a technical indicator that tracks stock momentum. The system is designed to be computationally efficient and accessible to individual investors, capable of running on average laptop GPUs, without the need for high-end hardware. Development and testing were conducted on a device with a 12th Gen Intel(R) Core(TM) i7-1260P processor operating at 2.10 GHz, 16 GB of RAM (15.7 GB usable), and a 64-bit operating system on an x64-based processor, demonstrating its compatibility with standard consumer-grade hardware. By combining seven days of news sentiment data with the RSI, the system provides actionable insights that are both timely and relevant to short-term market fluctuations. This approach helps reduce computational costs, while maintaining the ability to make informed, data-driven trading decisions.

The proposed system addresses the research problem by

- Using NLP for Sentiment Analysis: The system analyzes recent news articles to gauge market sentiment, providing traders with real-time insights into how news events might affect stock prices. This reduces reliance on historical data and allows traders to make decisions based on current market sentiment
- Combining Sentiment with RSI: By integrating sentiment analysis with RSI, the system provides a more comprehensive view of stock behavior, allowing traders to identify overbought or oversold conditions, alongside market sentiment.
- Improving Accessibility: The system is designed to be accessible to individual investors, running efficiently on average computing hardware, thus democratizing access to advanced trading tools that would otherwise require substantial investment.
- Offering Flexibility: Traders can personalize the system to fit their own strategies, enabling them to create a cost-effective and efficient solution for automated trading, without the high costs typically associated with such systems.

This system was tested on U.S. stocks across different market capitalizations (large, mid, and small caps), demonstrating its ability to adapt to various market conditions. Comparative analysis showed that the system is competitive in terms of accuracy, profitability, and computational efficiency. Ultimately, the system offers a practical, real-time trading strategy for individual investors by combining the strengths of NLP and technical analysis, enabling more accurate and accessible stock trading decisions. Our proposed system improves upon existing models by integrating real-time sentiment analysis with the RSI, offering a lightweight, context-aware, and computationally efficient approach to stock trading decisions, without compromising accuracy.

The rest of the paper is organized as follows: Section 2 provides a review of existing methods for stock market prediction, with a focus on NLP and technical analysis. Section 3

details the proposed solution, while Section 4 presents the experimental results, including a comparison with benchmark indices. Finally, Section 5 concludes the study and Section 6 highlights potential directions for future work in expanding the system's capabilities and addressing its limitations.

2. Related Works

Stock price prediction has been extensively studied using both sentiment analysis and technical indicators, each contributing unique advantages for forecasting stock movements. Sentiment analysis of financial news, blog posts, and investor discussions is widely used to determine the strength of market trends and predict future price movements [10–12]. Similarly, technical indicators such as moving averages (MA) and the RSI provide insights into market conditions by identifying overbought or oversold stocks, helping traders gauge trend strength and potential reversals [13]. Due to their complementary nature, researchers have explored integrating sentiment analysis with technical analysis to develop more effective stock recommendation systems.

Several recent studies have explored different combinations of these methodologies, each offering valuable insights into stock prediction, while also exhibiting certain limitations. Ref. [14] investigated the effectiveness of news sentiment analysis for investment strategies using a deep-learning-based sentiment analyzer. The study evaluated sentiment scores aggregated at daily, weekly, and monthly intervals and found that daily sentiment-based strategies yielded the highest returns but also required intensive short-term trading activity. This suggests that, while sentiment from financial news can be predictive, its impact on stock prices is often short-lived, particularly for high-profile news stories. However, a major limitation of this approach is its sole reliance on news sentiment, which ignores other market variables such as technical indicators, potentially reducing prediction accuracy.

Similarly, ref. [15] proposed a hybrid approach that combines historical stock data with financial news sentiment analysis. By integrating sentiment scores derived from multiple machine learning models, the study achieved up to 0.90 accuracy in both short- and long-term stock trend predictions. However, the study only evaluated four companies from different sectors, limiting its ability to generalize across the broader stock market. Additionally, it observed weaker predictive power for certain industries (e.g., the automobile sector), suggesting that sentiment's impact on stock prices may vary across different domains.

An alternative approach by [16] focused on integrating environmental, social, and governance (ESG) sentiment data with technical indicators for predicting stock prices, particularly for the S&P 500 index. The study employed bidirectional recurrent neural networks (Bi-RNN) and bidirectional long short-term memory networks (Bi-LSTM) to assess the influence of ESG sentiment on stock prices. The results demonstrated that incorporating ESG sentiment improves prediction accuracy when combined with technical indicators. However, the study was limited to S&P 500 stocks, making it difficult to generalize to other markets. Furthermore, ESG sentiment may have varying impacts, depending on the industry, necessitating more specialized industry-based models.

Other studies have also explored sentiment analysis approaches, with varying degrees of success. For instance, ref. [17] focused exclusively on market sentiment derived from investor reviews and financial news, without incorporating technical indicators. This limits the predictive accuracy, as it lacked a quantitative framework for evaluating market trends. Similarly, ref. [18] confirmed that financial news sentiment improves prediction performance but did not incorporate technical indicators, reducing its ability to detect price momentum and trend strength. Ref. [19] attempted to bridge this gap by combining news sentiment analysis with technical indicators such as RSI and SMA. However,

the study relied on a small dataset of only eight news articles, raising concerns about data representativeness and the robustness of sentiment analysis.

While existing methods have contributed significantly to stock prediction research, they suffered from key limitations, such as reliance on either sentiment analysis or technical indicators in isolation, limited datasets, or a lack of real-time adaptability. Our approach addresses these gaps by introducing a dual-layered analytical framework that integrates real-time sentiment analysis with technical indicators, particularly the RSI.

Unlike many previous studies that used small, static datasets, such as [19], which only considered eight articles for sentiment analysis, our system leverages up to 100 financial news articles in real time. This ensures a more comprehensive assessment of market sentiment, making the analysis more robust and representative of market conditions.

Many studies have relied exclusively on sentiment analysis, without incorporating technical indicators, as seen in [17,18]. While sentiment-based approaches can capture market sentiment fluctuations, they often fail to account for price momentum and trend strength. By integrating the RSI, our model compensates for this limitation, allowing us to capture both market sentiment and technical market conditions simultaneously.

Existing models, such as the ESG-focused approach by [16], have often been restricted to specific market subsets like the S&P 500. In contrast, our method is not industry-specific and can be applied to various stocks across different sectors, increasing its adaptability. Additionally, sentiment-based models that rely solely on daily news impacts, such as [14], often struggle with short-lived predictive power. By incorporating technical analysis, our hybrid model extends the predictive stability of sentiment-based signals, ensuring more stable and actionable trading recommendations.

By combining qualitative (news sentiment) and quantitative (RSI) factors, our approach outperforms traditional single-layered models and provides a more holistic perspective on stock price movements, as shown in Table 1. This significantly improves prediction accuracy, investment performance, and adaptability across diverse market conditions. Our work extends the capabilities of existing stock prediction systems by integrating real-time sentiment analysis with technical indicators, creating a more comprehensive and adaptive framework. Unlike prior works that relied on small datasets, our system leverages up to 100 news articles, for a more robust sentiment assessment. By bridging the gap between sentiment impact and trend analysis, our dual-layered approach enhances predictive accuracy and investment performance. Furthermore, while existing sentiment analysis-based models have shown effectiveness, they often overlook price momentum, which we address by incorporating the RSI. The flexibility of our model allows it to be applied across different industries and market conditions, unlike sector-specific approaches such as ESG sentiment analysis. This adaptability, coupled with a more stable prediction framework, positions our model as an advancement over previous stock prediction techniques. By addressing the shortcomings of earlier works and enhancing the synergy between sentiment and technical analysis, our approach provides a robust, scalable, and high-precision stock prediction system, with the potential to significantly improve investment decision-making.

Table 1. Comparison of stock prediction approaches, including our proposed model.

Study	System/Technique	Sentiment Analysis Approach	Technical Indicators Used	Key Difference/Contribution
[10–12]	Sentiment Analysis (News, Blogs)	News-based sentiment analysis	None	<ul style="list-style-type: none"> - Identifies bullish/bearish signals in the stock market - Useful for forecasting future price movements - Limited effectiveness, due to short lifespan of sentiment data

Table 1. Cont.

Study	System/Technique	Sentiment Analysis Approach	Technical Indicators Used	Key Difference/Contribution
[13]	Technical Analysis (MA, RSI)	None	MA, RSI	<ul style="list-style-type: none"> - Diagnoses overbought/oversold conditions - Detects trend strength for trading timing - Lacks external sentiment-based data
[14]	Deep-Learning-Based News Sentiment Analysis	Sentiment analysis on financial news	None	<ul style="list-style-type: none"> - High returns with daily sentiment data - High turnover required with daily strategies - Excludes technical indicators
[15]	Integrated Model with News Sentiment + Historical Data	News sentiment analysis	Historical stock data (not specified)	<ul style="list-style-type: none"> - Achieved high accuracy (up to 0.90) for forecasting - Limited generalizability, due to a small number of companies
[16]	ESG Sentiment + Technical Indicators	ESG sentiment index from news	Multiple indicators	<ul style="list-style-type: none"> - ESG sentiment improves prediction performance - Limited to S&P 500 dataset
[17]	Market Sentiment Analysis	Investor reviews and news-based sentiment	Not specified	<ul style="list-style-type: none"> - Focuses on sentiment analysis without technical indicators like RSI
[18]	Financial News Sentiment Model	Sentiment analysis on financial news	Not specified	<ul style="list-style-type: none"> - Similar sentiment analysis approach but does not incorporate RSI
[19]	Sentiment Analysis + Technical Indicators	Sentiment analysis based on news	RSI, SMA	<ul style="list-style-type: none"> - Integrates RSI with technical indicators but lacks large-scale sentiment analysis - Used a small dataset of only 8 articles for stock prediction
Our Approach	Hybrid Sentiment + Technical Model	Real-time news sentiment analysis	RSI for trend detection	<ul style="list-style-type: none"> - Integrates the RSI with real-time sentiment analysis - Dual-layered, multi-dimensional approach - Uses up to 100 news articles for sentiment evaluation, improving robustness

3. Materials and Methods

We aimed to integrate sentiment analysis with conventional technical analysis, to provide buy-and-sell recommendations in the stock market. We specifically propose an algorithm that combines sentiment and technical analysis, to make buy-and-sell decisions for stocks. The proposed model aims to assist investors in making well-timed decisions by increasing the likelihood of buying or selling stocks at the optimal moment. Our system classifies trading signals into categories such as strong buy, buy, neutral, sell, and strong sell. These classifications are based not only on financial data, including open price, close price, high price, low price, and volume, but also on sentiments extracted from corporate news. The conceptual framework of the model is illustrated in Figure 1.

3.1. Input

Initially, the system receives a ticker (stock symbol) as input. A ticker is a unique series of letters assigned to a security for trading purposes [20]. Stocks listed on the New York Stock Exchange (NYSE) can have four or fewer letters. Nasdaq-listed securities can have up to five characters. For example, “AAPL” for Apple or “NVAX” for Novavax. In addition, some stocks trade on the over-the-counter markets (OTCMKTS), which are decentralized and not listed on formal exchanges like the NYSE or Nasdaq. OTC tickers often indicate smaller or foreign companies and can vary in format. An example is “TSNP” for HUMBL Inc., San Diego CA, USA.

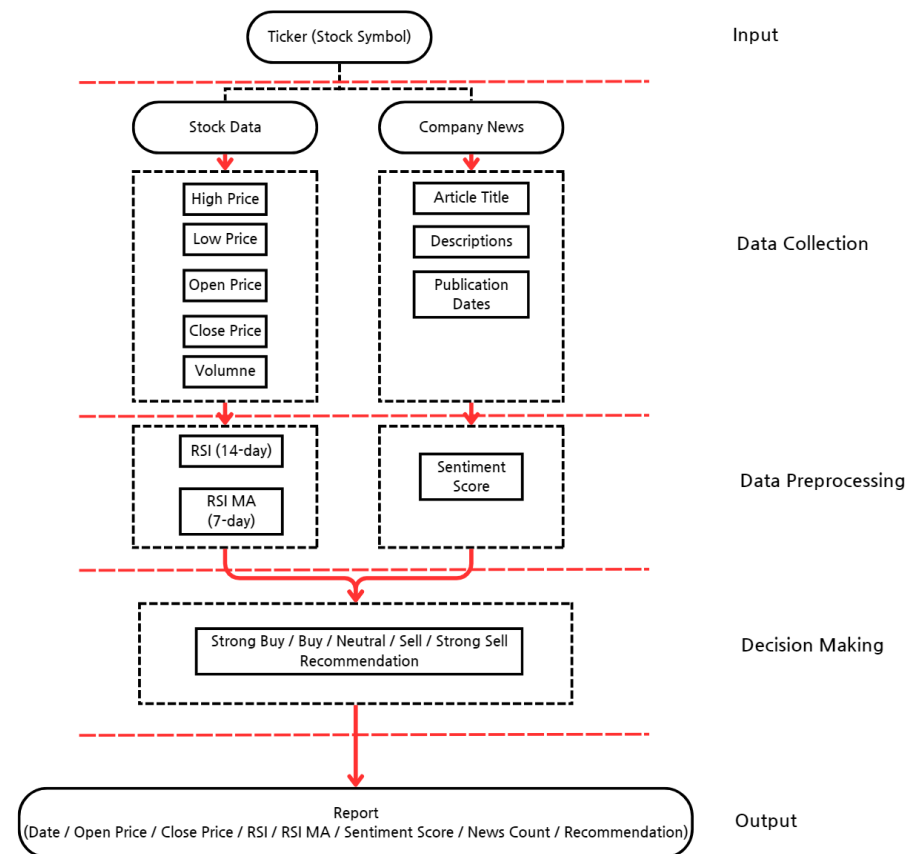


Figure 1. A conceptual framework for a rule-based stock trading recommendation system incorporating sentiment analysis and technical indicators.

3.2. Data Collection

Once the inputs have been provided, the system gathers two types of data: 1. stock data and 2. news data. For stock data, the system retrieves historical information. For stock data, the system retrieves daily opening and closing prices for the past two months, focusing on the most recent 30 trading days to generate detailed reports. In this experiment, we obtained U.S. stock data from Polygon.io (accessed on 9 January 2025) [21], using it as the source for stock prices.

For news data, the system collects relevant articles (titles, descriptions, and publication dates) from the past 7 days for each trading day. For example, for August 8, the system gathers all news published between August 1 and August 8. In this experiment, we used a news API to fetch news articles for the sentiment analysis. The API request was structured as follows:

- API Endpoint: (accessed on 9 January 2025) <https://newsapi.org/v2/everything> [22]
- API Parameters:
 - q: Stock ticker symbol (e.g., “AAPL” for Apple)
 - from: Start date for fetching news
 - to: End date for fetching news
 - sortBy: Sort by publication date
 - language: Only fetch English news (en)

The output of the NewsAPI, as shown in Listing 1, is structured in JSON format and contains three primary components: status, totalResults, and articles. The status field indicates whether the API request was successful (e.g., “ok”), while totalResults provides the total number of articles matching the query criteria.

The articles field is an array of objects, where each object represents a single news article. Each article contains key-value pairs, such as

- source: Specifies the news source, including its id (if available) and name.
- author: The author of the article.
- description: A brief summary of the article.
- url: The link to the full article.
- urlToImage: The URL for an image associated with the article (if available).
- publishedAt: The publication date and time in ISO 8601 standard [23].
- content: The main content or excerpt of the article.

This hierarchical structure ensures that detailed information about each news item is easily accessible and programmatically retrievable for analysis.

Listing 1. Sample JSON Output from NewsAPI.

```
{
  "status": "ok",
  "totalResults": 2072,
  "articles": [
    {
      "source": {
        "id": "the-verge",
        "name": "The Verge"
      },
      "author": "Andrew Liszewski",
      "title": "Apple's Sports app now tells you where you can watch
        nationally broadcast games",
      "description": "Apple's Sports app has been updated with broadcast
        information for games and improved navigation letting you
        quickly swipe between teams and league pages.",
      "url": "https://www.theverge.com/2025/1/27/24353035/apple-sports-
        app-ios-iphone-broadcast-information-where-to-watch",
      "urlToImage": "https://cdn.vox-cdn.com/thumbor/
        ssk31YCSGHejHI0bwmcbgqz6pSE=/cdn.vox-cdn.com/uploads/
        chorus_asset/file/25848941/apple_sports2.jpg",
      "publishedAt": "2025-01-27T17:32:12Z",
      "content": "Apple's Sports app now tells you where you can watch
        nationally broadcast games..."
    }
  ]
}
```

To ensure the reliability and objectivity of the proposed system, this study relies exclusively on news data as the primary source of sentiment analysis. While social media sentiment can provide valuable insights, due to its pervasive influence on public opinion and market dynamics, it is often highly subjective and emotionally driven [24,25]. This subjectivity can introduce significant noise into the analysis, potentially affecting the accuracy of the predictions [24]. By contrast, news articles are typically curated and adhere to professional journalistic standards, offering a more consistent and factual representation of market events [25]. This focus on news sentiment aligns with the system's design philosophy of operating on information that is as objective and reliable as possible.

3.3. Data Preprocessing

The data are processed through two parallel analyses: 1. technical stock analysis and 2. sentiment analysis. Stock investment strategies can generally be divided into two major approaches: technical analysis and fundamental or industry analysis [26]. These two approaches complement each other, providing a comprehensive framework for making informed investment decisions [27]. While technical analysis focuses on predicting future price movements based on historical price patterns, industry analysis seeks to interpret external factors and market sentiment to understand the underlying drivers of price fluctuations. We combined the quantitative foundation of technical analysis with the qualitative insights of news sentiment analysis, effectively addressing the limitations both methods when used independently.

3.3.1. Technical Stock Analysis

The proposed algorithm leverages one of the key components of technical analysis by incorporating the relative strength index (RSI) as its primary indicator. The relative strength index (RSI) is defined as follows:

$$RSI = 100 - \left(\frac{100}{1 + RS} \right) \quad (1)$$

where

$$RS = \frac{\text{Average Gain over } n \text{ periods}}{\text{Average Loss over } n \text{ periods}} \quad (2)$$

and n is typically set to 14 periods by default.

In the RS formula, the average gain is derived from the average of upward price changes over a specified period, while the average loss represents the average of downward price changes over the same period. Notably, the average loss is expressed as a positive number, ensuring that RS is always greater than or equal to zero. As a result, RS ranges from zero to infinity, while the RSI remains bounded within the range of 0 to 100 [28].

An RSI value above 70 typically indicates an overbought condition, signaling a potential downturn, whereas a value below 30 suggests an oversold condition, indicating a potential upward reversal. These thresholds help traders identify potential entry and exit points in the market. The RSI's bounded nature makes it particularly useful for comparing the relative momentum between different stocks or assets. Furthermore, its reliance on both gains and losses offers a balanced perspective, avoiding the bias inherent in analyzing price movements in only one direction.

3.3.2. Sentiment Analysis

While the RSI effectively identifies trading opportunities based on price trends, it cannot explain the root causes behind price movements [29]. To address this limitation, the algorithm incorporates news sentiment analysis to account for external market influences and investor sentiment. The role of sentiment analysis in the algorithm involves evaluating the emotional tone of market-related textual data, such as news articles. When the news shows a predominance of positive sentiments, this indicates optimism in the market, which may signal an upcoming upward trend. Conversely, if negative sentiment dominates, this reflects pessimism and could suggest potential downward pressure on the market [30].

For sentiment analysis, we use TextBlob [31], a natural language processing tool, to assess the polarity of news articles. Polarity values range from -1 (most negative) to 1 (most positive), and for each 7-day period, the sentiment scores are averaged and scaled between 1 and 100. This scaling ensures that a value of 50 is neutral, with higher values

indicating positive sentiments and lower values indicating negative sentiments. TextBlob processes stopwords by removing them during tasks such as tokenization and text cleaning. Stopwords, such as “the”, “is”, “in”, and “and”, are common words that contribute little meaning to the analysis. The tool uses a predefined stopwords list from libraries like NLTK to filter out these terms. While TextBlob does not specifically handle abbreviations unless additional preprocessing steps are added, this was not necessary for our sentiment analysis of news articles, since formal texts like news articles typically have fewer abbreviations [32].

- Polarity: Ranges from -1 (most negative) to 1 (most positive).
- Formula for Average Sentiment Score:

$$\text{Average Sentiment} = \frac{\sum \text{Polarity of articles}}{\text{Number of articles}}$$

- Sentiment Score Scaling: The system scales the average sentiment score to a value between 1 and 100 for easier interpretation:

$$\text{Scaled Sentiment} = (\text{Average Sentiment} + 1) \times 50$$

3.4. Decision-Making

Based on the RSI and sentiment score, the system generates buy–sell recommendations for the 15 most recent trading days. The logic for these decisions is as follows:

- Strong Buy: If the sentiment score is above 70 (highly positive) and the RSI is below 30 (indicating oversold conditions), the system issues a “Strong Buy” recommendation.
- Buy: If the sentiment score is above 50 (moderately positive) and the RSI falls between 30 and 50, the system issues a “Buy” recommendation.
- Neutral: If the sentiment score is between 40 and 60, and the RSI also ranges between 40 and 60, the system issues a “Neutral” recommendation.
- Sell: If the sentiment score is below 50 (moderately negative) and the RSI is above 70 (indicating overbought conditions), the system issues a “Sell” recommendation.
- Strong Sell: If the sentiment score is below 30 (highly negative) and the RSI is above 70, the system issues a “Strong Sell” recommendation.

The algorithm’s use of the RSI is guided by two foundational theories. The first is Dow theory, which divides market movements into six distinct phases: accumulation, uptrend, overheating, distribution, panic, and downtrend. The algorithm applies the RSI to identify these phases and adjusts buy–sell recommendations accordingly [33]. For example, during the overheating phase, the RSI signals indicate selling, while during the panic or downtrend phases, the RSI signals suggest buying. The second is Granville’s investment decision, which emphasizes that buying is most effective during panic or downtrend phases, while selling is optimal during uptrend or overheating phases [34]. This aligns closely with the RSI’s ability to identify overbought and oversold conditions, providing clear guidance for the algorithm’s recommendations.

The complementarity of news sentiment analysis adds a qualitative layer to the quantitative insights provided by the RSI. For example, if the RSI signals an oversold condition, but the prevailing sentiment in the news is highly positive, the algorithm considers the possibility of a potential upward reversal, suggesting a buy signal despite the oversold condition. Conversely, if the RSI indicates an overbought condition, but the news sentiment is highly negative, the algorithm accounts for the potential for downward pressure and issues a sell recommendation. This integration ensures a more nuanced understanding of market dynamics, reducing reliance on solely historical price patterns.

The combined approach of integrating the RSI and news sentiment analysis allows the algorithm to leverage both quantitative and qualitative data streams, creating a multi-dimensional approach to stock recommendations. Quantitative insights from the RSI offer precise buy–sell signals based on market trends and momentum. Qualitative insights from sentiment analysis refine these recommendations by interpreting the emotional and contextual factors behind market movements.

3.5. Output

For each of the last 30 trading days, the system generates a detailed report that includes several key data points, as shown in Table 2.

Table 2. Key data elements in the report.

Data Elements	
Date	The specific trading day being analyzed.
Open Price	The opening price of the stock on that day.
Close Price	The closing price of the stock on that day.
High Price	The highest price of the stock on that day.
RSI	The 14-day RSI for that day.
RSI Moving Average (RSI MA)	The 7-day moving average of the RSI.
Sentiment Score	The overall sentiment score based on news articles from the preceding 7 days.
Number of News Articles	The number of news articles used for sentiment analysis.
Recommendation	The system’s recommendation for that day, whether Buy, Sell, Strong Buy, Strong Sell, or Neutral.

The report is presented in a clear, human-readable format, summarizing all the key data points for easy interpretation by the user, as shown in Listing 2. The Date represents the specific trading day under review. The Open Price and Close Price indicate the stock’s price at the market’s opening and closing, respectively, while the High Price reflects the highest price reached during the day. The RSI is the 14-day relative strength index, a technical indicator used to identify overbought or oversold conditions. The RSI MA is the 7-day moving average of the RSI, smoothing out fluctuations for better trend analysis. The Sentiment Score is derived from analyzing news articles from the previous seven days, providing an overall market sentiment ranging from negative to positive. The Number of News Articles indicates the volume of news data used for sentiment analysis. Finally, the Recommendation reflects the system’s trading suggestion for that day, such as Buy, Sell, Strong Buy, Strong Sell, or Neutral, based on the combined analysis of technical and sentiment indicators.

Listing 2. System Input and Output for BABA.

```
$ python script_name.py~BABA

Date: 2025-01-17, Open: 83.20, High: 85.59, Close: 85.12, RSI:
50.79, RSI MA: 39.69, Sentiment Score: 54, News Count: 99,
Recommendation: Neutral
Date: 2025-01-16, Open: 82.02, High: 82.59, Close: 82.43, RSI:
40.97, RSI MA: 38.56, Sentiment Score: 54, News Count: 99,
Recommendation: Buy
```

4. Results

To thoroughly evaluate the effectiveness of our proposed model, we designed and conducted experiments aimed at providing a comprehensive assessment of its performance. These experiments were structured to validate the model's predictive accuracy and profitability across a diverse set of conditions, including different stock subsets and multiple time periods.

Our initial validation was conducted on a dataset comprising 95 U.S. stocks categorized into large-cap, mid-cap, and small-cap groups. This backtest spanned a seven-week period, from 11 September 2024, to 31 October 2024. While these results provided valuable insights, the scope of the initial validation was limited to a single period and stock categorization.

To address these limitations and strengthen our analysis, we expanded the experimental framework to include multiple time periods. Specifically, the model's performance was tested over two distinct intervals: 2 weeks and 1 month. These additional time frames allowed for a more robust evaluation, capturing both short-term and medium-term market behaviors. Furthermore, stocks were categorized based on market characteristics, such as beta (high beta, medium beta, and negative beta). This comprehensive approach ensured a deeper understanding of the model's adaptability across various subsets of stocks and time horizons.

4.1. Comparative Rate of Return Analysis

To validate our model's predictive accuracy and profitability, we conducted a backtest on a diverse set of 95 U.S. stocks, categorized by large-cap, mid-cap, and small-cap groups, as shown in Table 3. This backtest covered a seven-week period from 11 September 2024, to 31 October 2024, using data sourced from [Polygon.io](#) [21] for stock prices and [NewsAPI.org](#) [22] to gather relevant news.

Table 3. List of tickers used in the experiment by market cap size.

Market Cap Size	Tickers
Large Cap (34)	APPL, MSFT, NVDA, GOOGL, AMZN, META, LLY, TSM, AVGO, JPM, UNH, NVO, TCEHY, XOM, PG, ABBV, ASML, IDCBY, MRK, CVX, AMD, CRM, RHHBY, HESAY, TMUS, AZN, TM, LRLCY, TMO, PEP, ACGBY, ADBE, NVS, ACN
Mid Cap (30)	QGEN, KEP, BWXT, SNX, ONTO, CHDN, ULS, BSAC, XP, FRT, MTCH, OLED, AUR, GME, YMM, EVR, MTZ, SF, QRVO, GMED, GNRC, EXP, APA, TPX, BIO, LW, RGLD, GL, CX, STN
Small Cap (31)	BANR, VCEL, VSCO, VYX, MRVI, HC, SIMO, KMT, MIRM, MBIN, FLYW, OFG, CABO, NATL, TALO, OMCL, PINC, BEAM, AGM, ADNT, SLNO, ASPN, SAH, SDHC, AG, WSBC, CEPU, HI, PLUG, MODG, TRMK

To ensure that news data up to the previous day are reflected in our system, we designed it to execute purchases at the next day's opening price. This approach is based on the observation that significant trading activity often occurs at the opening price following major news releases, resulting in higher trading volumes during this period. By targeting the opening price, we aim to capitalize on this increased liquidity, thereby reducing the likelihood of unfilled orders. Research indicates that after-hours trading and overnight news can significantly influence the opening price of a stock on the following day, leading to notable price movements at market opening [35].

To evaluate our system, we simulated trades based on the recommendations generated by the system, employing a structured trading algorithm, as shown in Algorithm 1. The algorithm dynamically adjusts positions to maximize returns, while managing risk.

Algorithm 1 Trading algorithm

```

1: Start
2: if Signal is BUY or Strong BUY then
3:   Buy at the next day's open price.
4:   while Signal remains BUY or Strong BUY do
5:     if Open price is below the last buy price then
6:       Buy additional shares.
7:     end if
8:   end while
9: end if
10: if Signal changes to Neutral or Lower then
11:   if Net profit  $\geq 1\%$  then
12:     Sell at the open price.
13:   else if Holding period  $\geq 14$  days then
14:     if Net profit  $\geq 0.25\%$  then
15:       Sell at the open price.
16:     end if
17:   end if
18: end if
19: End

```

Net Profit: The net percentage return from a trade, calculated as

$$\text{Net Profit} = \frac{(\text{Selling Price} - \text{Average Purchase Price})}{\text{Average Purchase Price}} \times 100 \quad (3)$$

Transaction costs, such as trading fees and potential slippage, are not accounted for in this calculation.

Holding Period: The number of trading days since the initial purchase of a stock. If a position is held for an extended period without reaching the primary profit target, a lower threshold is applied to exit the trade.

The trading strategy follows a systematic approach to buying and selling stocks based on sentiment signals and technical indicators. The trading rules are defined as follows:

- **Buy Strategy:** When a “BUY” or “Strong BUY” signal appears, a purchase is made at the next day’s open price. If the signal remains at “BUY” or “Strong BUY” in subsequent days, additional purchases are only made when the open price falls below the most recent purchase price.
- **Sell Strategy:** When a “Neutral” or lower signal (Neutral, Sell, Strong Sell) appears, the stock is sold at the open price. However, for “Neutral” signals, a sale is only executed if the net profit exceeds 1%.
- **Holding Period Adjustment:** If a stock is held for 14 days without reaching the 1% target, the minimum profit threshold is adjusted to 0.25%, allowing a sale if the stock gains at least 0.25% after the 14-day holding period.

The algorithm is designed to optimize trading performance by balancing profit-taking with risk management. When a **BUY** or **Strong BUY** signal is triggered, an initial purchase is made, with provisions for additional purchases under favorable conditions. The exit strategy prioritizes securing profits, while preventing excessive holding periods. A stock is sold once a **Neutral** or lower signal appears, provided that the profit threshold conditions are met. The introduction of the 14-day holding rule ensures that capital is not locked into unproductive positions, enabling a more dynamic trading approach.

This trading algorithm allowed us to measure performance in line with the system’s recommendations, accurately reflecting the potential profitability and effectiveness of the model’s signals. The rate of return (RoR) is a fundamental metric used to measure the

performance of an investment over a specific period. It represents the percentage change in the value of the investment, providing a standardized way to evaluate profitability or loss. The formula for calculating the rate of return is as follows:

$$\text{Rate of Return (RoR)} = \frac{\text{Final Value of Investment} - \text{Initial Value of Investment}}{\text{Initial Value of Investment}} \times 100 \quad (4)$$

where

- **Final Value of Investment:** The value of the investment at the end of the specified period, including any capital appreciation and realized gains.
- **Initial Value of Investment:** The value of the investment at the beginning of the period.
- The result is expressed as a percentage.

To provide an objective evaluation, we compared the performance of our model against widely recognized benchmark indices, which are crucial for evaluating investment strategies. Benchmark indices are not only formal measures of investment performance but also serve as critical tools in constructing and assessing investment portfolios [36]. For this purpose, we used prominent U.S. market indices, including the NASDAQ Composite Index (COMP), the Dow Jones Industrial Average (DJIA), and the S&P 500 Index. The NASDAQ Composite Index is ideal for assessing performance in technology-driven and high-growth sectors. The DJIA focuses on large-cap stocks, offering a measure of stability and market leadership, while the S&P 500 provides a comprehensive representation of the overall U.S. equity market. Comparing our model against these indices offered an objective standard to evaluate its profitability and predictive effectiveness.

An average return of 4.03% was achieved across 95 stocks, with a 5.10% increase compared to the initial investment capital, based on a total purchase amount of \$22,222.09 and a total sale amount of \$23,355.92. This performance was also compared against the COMP, DJIA, and S&P 500 indices, as shown in Table 4. The annualized rate of return for our model over the period from 11 September 2024 to 31 October 2024 was 5.10%, compared to 2.21%, 4.02%, and 2.72% for the DJIA, COMP, and S&P 500, respectively.

Table 4. Rate of return comparison with major market indices.

Metric	DJIA	COMP	S&P 500	Our Model
RoR (%)	2.21	4.02	2.72	5.10

The RoR analysis presented in this study was based on the assumption that transactions are executed using limit orders, which are designed to buy or sell stocks at a specified price or better. This approach minimizes the impact of market slippage, as trades are only executed under predefined price conditions. Consequently, market slippage is expected to have a negligible effect on the system's performance. However, we acknowledge that transaction costs, such as brokerage fees, and risk-adjusted metrics, such as the Sharpe ratio, are essential for providing a more comprehensive evaluation of the system. These factors will be incorporated in future work, to enhance the robustness of the analysis and offer a deeper understanding of the system's real-world applicability.

This analysis demonstrates the model's capability to generate returns that outperform these benchmarks, underscoring its potential as a profitable investment tool.

4.2. Expanded Experimental Results

We conducted additional experiments to provide a more comprehensive evaluation of the models predictive accuracy and profitability. These experiments were designed to

test our approach across different subsets of stocks and multiple time periods, ensuring a robust assessment of its performance under various market conditions.

To evaluate the adaptability of our model, stocks were categorized based on two criteria.

First, stocks were classified based on their market capitalization, a widely recognized metric that categorizes companies by their total market value. This classification grouped stocks into three categories: Large Cap, Mid Cap, and Small Cap, facilitating a comprehensive evaluation of the model's performance across different market segments. Large-cap stocks represent companies with the highest market capitalization, often associated with stability and lower volatility, due to their established market presence and significant resources [37]. These firms generally include industry leaders with substantial market share. In contrast, mid-cap stocks comprise companies with moderate market capitalization. These businesses are typically in a growth phase, offering a balance between the stability of large-cap stocks and the growth potential of smaller firms, though they may carry higher risks [38]. Small-cap stocks, characterized by the smallest market capitalizations, are often associated with emerging companies or niche market players. These stocks exhibit higher volatility but also present substantial growth potential, appealing to investors seeking significant returns in exchange for elevated risks [39]. This cap-size-based categorization ensured a thorough evaluation of our model's performance across diverse market segments. It captured the varying risk–return profiles and behavioral dynamics characteristic of different capitalization tiers. By analyzing results within these categories, the study aimed to identify patterns and trends that enhance the applicability and robustness of the model in distinct market contexts.

Second, stocks were grouped by beta size into high-beta ($\beta \geq 1$), medium-beta ($0 < \beta < 1$), and negative-beta ($\beta \leq 0$) categories. In finance, the *beta coefficient* (β) measures the sensitivity of a security's returns to movements in the overall market. It quantifies the extent to which a particular asset's price fluctuates in response to market changes, serving as an indicator of the asset's systematic risk relative to the market [40]. Mathematically, beta is calculated through regression analysis, where the returns of the security are regressed against the returns of a market benchmark, such as the S&P 500 Index. The formula for beta is given as

$$\beta = \frac{\text{Cov}(R_i, R_m)}{\text{Var}(R_m)}$$

where

- $\text{Cov}(R_i, R_m)$ represents the covariance between the returns of the security (R_i) and the market returns (R_m).
- $\text{Var}(R_m)$ denotes the variance of the market returns.

Covariance measures how two variables move together, while variance quantifies the dispersion of a set of numbers. Beta, derived from these metrics, measures a security's sensitivity to market movements and can be classified into three categories:

1. **Beta** ≤ 0 : Securities with a beta less than or equal to 0 exhibit an inverse relationship with the market.
2. $0 < \text{Beta} < 1$: Securities in this range demonstrate lower volatility relative to the market.
3. **Beta** ≥ 1 : Securities with a beta greater than or equal to 1 exhibit higher volatility compared to the market.

This classification framework facilitates a nuanced analysis of performance across varying levels of market volatility and sensitivity, enabling informed investment decisions.

The time periods selected for backtesting included two weeks and one month, with the end date set as 24 January 2025. Selecting appropriate time frames for backtesting is crucial for evaluating the effectiveness of trading strategies. For short-to-midterm trading strategies, which involve holding positions for a few days to several weeks, backtesting over periods of several weeks to a few months is generally considered sufficient [41]. This duration allows traders to assess how their strategies perform across various market conditions, without overfitting to a specific time frame. By focusing on these two distinct periods, the analysis captured the performance of the model under different temporal conditions, while maintaining a practical testing framework.

To evaluate the performance of our stock recommendation algorithm, we utilized the following metrics:

- **Mean (Average Return):** The central tendency of the rates of return (RoR) across the selected stocks or groups.
- **Median:** The midpoint of the RoR data, indicating the typical return, while minimizing the influence of outliers.
- **Standard Deviation:** A measure of variability, reflecting the dispersion of RoR values around the mean.
- **Minimum and Maximum:** The lowest and highest returns observed, highlighting the extremes in performance.
- **95% Confidence Interval:** The range within which the true mean RoR is expected to lie, with 95% certainty.
- **Coefficient of Variation (CV):** The ratio of standard deviation to the mean, expressed as a percentage, indicating the relative variability in returns.
- **Skewness:** A metric of asymmetry in the return distribution, indicating potential biases toward gains or losses.
- **Kurtosis:** A measure of the tailedness of the RoR distribution, identifying the prevalence of extreme outcomes.
- **Accuracy:** The accuracy metric is calculated as the percentage of predictions, where the predicted direction of the rate of return (positive or negative) matches the actual direction.

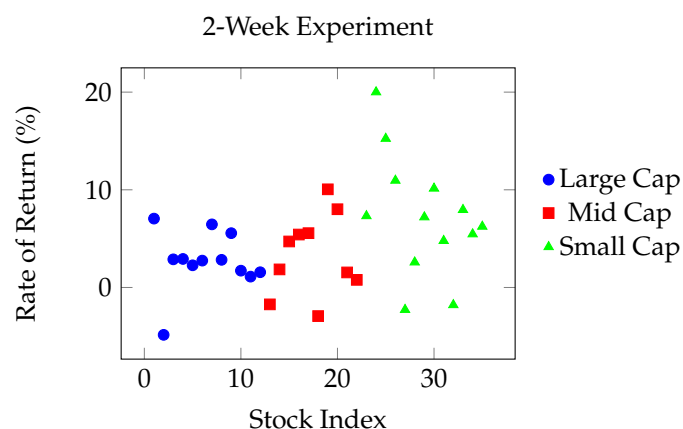
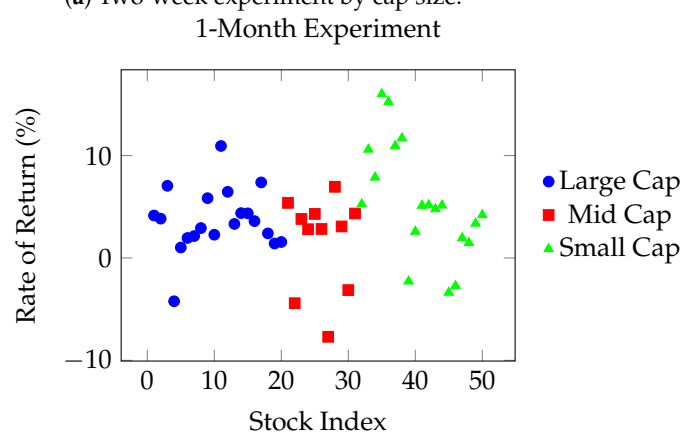
In addition to analyzing these metrics, we further benchmarked the algorithm's performance against relevant market indices. By doing so, we assessed not only the algorithm's standalone performance but also its relative effectiveness compared to broader market trends.

4.2.1. Performance Analysis by Cap Size

The cap size experiment aimed to evaluate the stock trading recommendation system's performance across three categories of stocks: Large Cap, Mid Cap, and Small Cap. Table 5 lists the stock tickers grouped by their market capitalization, providing the basis for this analysis. Figure 2a,b display the scatter plots of RoR for the 2-week and 1-month experimental periods, respectively. These visualizations highlight the distribution of returns and help identify trends in each cap size group. Additionally, Tables 6 and 7 present detailed summary statistics, including mean, median, standard deviation, and accuracy metrics for each group. Lastly, Table 8 compares the system's performance with benchmark indices.

Table 5. Ticker grouping by cap size.

Cap Size	Tickers
Large Cap	NVDA, AAPL, MSFT, AMZN, GOOG, META, TSLA, AVGO, TSM, WMT
Mid Cap	NICE, LKQ, BIO, RBC, WPP, UWMC, NBIS, BRBR, MGM, DBX
Small Cap	INTR, APLD, BTE, FINV, OFG, ATRC, FL, HSAI, AGIO, PRM

**(a)** Two-week experiment by cap size.**(b)** One-month experiment by cap size.**Figure 2.** Side-by-side comparison of RoR scatter plots for 2-week and 1-month experiments by cap size.**Table 6.** Detailed summary statistics and accuracy for cap size groups (2-week testing results).

Group	Mean	Median	Std Dev	Min	Max	95% CI	CV (%)	Skewness	Kurtosis	Accuracy (%)
Overall	4.76	4.71	4.65	−4.85	20.00	[3.14, 6.37]	97.69	0.68	−0.32	85.71
Large Cap	3.28	2.88	2.05	−4.85	7.05	[2.00, 4.56]	62.50	0.50	−0.83	83.33
Mid Cap	3.64	2.63	4.38	−2.93	10.05	[1.00, 6.27]	120.33	0.51	−0.88	80.00
Small Cap	7.36	7.95	6.38	−2.28	20.00	[4.06, 10.65]	86.70	0.67	−0.29	91.67

Table 7. Detailed summary statistics and accuracy for cap size groups (1-month testing results).

Group	Mean	Median	Std Dev	Min	Max	95% CI	CV (%)	Skewness	Kurtosis	Accuracy (%)
Overall	4.45	3.84	4.64	−7.70	16.01	[2.93, 5.97]	104.27	0.68	−0.28	81.63
Large Cap	4.16	3.32	3.12	−4.23	10.94	[2.86, 5.45]	75.00	0.56	−0.91	85.00
Mid Cap	2.27	3.08	4.32	−7.70	6.95	[0.01, 4.54]	190.31	−0.85	1.19	66.67
Small Cap	6.91	5.26	6.07	−3.38	16.01	[4.28, 9.54]	87.83	0.81	−0.12	90.00

The overall mean RoR for the 2-week period was 4.76%, with Small Cap stocks achieving the highest mean RoR of 7.36%. Large Cap and Mid Cap stocks exhibited mean RoRs of 3.28% and 3.64%, respectively. These results align with the higher volatility and growth potential typically associated with Small Cap stocks, which often yield higher returns during favorable market conditions. In contrast, Large Cap stocks showed more stability, reflected in their lower mean RoR but tighter clustering in the scatter plots. Mid Cap stocks, positioned between these two extremes, exhibited moderate performance. For the 1-month period, the overall mean RoR slightly decreased to 4.45%. Small Cap stocks continued to lead, with a mean RoR of 6.91%, while Large Cap stocks demonstrated consistent performance, with a mean RoR of 4.16%. Mid Cap stocks underperformed during this period, with a mean RoR of only 2.27%. The scatter plots for this period illustrate the greater dispersion in returns for Small Cap stocks, indicative of their high-risk, high-reward nature. Large Cap stocks showed tighter clustering, reinforcing their reputation for stability and reliability. The system's accuracy across cap sizes varied significantly. During the 2-week period, Small Cap stocks achieved the highest accuracy at 91.67%, followed by Large Cap stocks at 83.33%, and Mid Cap stocks at 80.00%. Over the 1-month period, the accuracy for Small Cap stocks decreased slightly to 90.00%, while Large Cap stocks improved to 85.00%. Mid Cap stocks had the lowest accuracy across both periods, with a 66.67% accuracy in the 1-month period, suggesting potential areas for refinement in the system's predictive capabilities for this category.

Table 8. Comparison of mean RoR for cap size groups and benchmarks (2-week and 1-month testing results).

Category	Mean RoR (2-Week, %)	Mean RoR (1-Month, %)
Overall	4.76	4.45
Large Cap	3.28	4.16
Mid Cap	3.64	2.27
Small Cap	7.36	6.91
COMP	3.49	1.36
DJIA	4.79	0.40
S&P 500	3.97	1.10

When comparing the system's performance to benchmark indices, the results were consistently favorable. During the 2-week period, the system's mean RoR of 4.76% exceeded the performance of the COMP (3.49%), DJIA (4.79%), and S&P 500 (3.97%). For the 1-month period, the system's overall mean RoR of 4.45% significantly outperformed the COMP (1.36%), DJIA (0.40%), and S&P 500 (1.10%). These results underscore the system's ability to generate superior returns compared to traditional market indices, regardless of the time frame.

The findings demonstrate that our stock trading recommendation system performed effectively across all cap size categories, with particularly strong results for Small Cap stocks. The system consistently outperformed benchmark indices in both the 2-week and 1-month periods, highlighting its robustness in diverse market conditions. The observed patterns align with the typical risk–return profiles of different cap sizes, where Small Cap stocks provide the highest potential for returns but exhibit greater variability, while Large Cap stocks deliver stable and reliable performance. However, the relatively low performance and accuracy for Mid Cap stocks suggest that further optimization may be needed to improve predictions for this category. Overall, the results affirm the system's ability to deliver strong investment outcomes across varying market segments.

4.2.2. Performance Analysis by Beta Sensitivity

The results from the two-week testing period highlight important trends and implications regarding the model's effectiveness and the characteristics of each beta group. The stock data used for this analysis comprised a diverse selection of 30 U.S. stocks, evenly distributed across three beta categories: high beta ($\beta \geq 1$), medium beta ($0 \leq \beta < 1$), and negative beta ($\beta < 0$), as shown in Table 9. Stocks in each category were carefully chosen to represent varying levels of market volatility and responsiveness. High beta stocks included well-known companies like NVIDIA (NVDA), Tesla (TSLA), and Amazon (AMZN), which are typically more volatile and sensitive to market trends. Medium beta stocks included companies such as Microsoft (MSFT), Walmart (WMT), and Procter & Gamble (PG), known for their stability and moderate correlation with the market. Negative beta stocks included Zoom (ZM), Biogen (BIIB), and TAL Education Group (TAL), which exhibit inverse relationships with the market, often acting as hedges during downturns. This categorization ensured a robust evaluation of the model's performance across varying market dynamics and stock characteristics.

Table 9. Ticker grouping by beta size.

Beta Category	Tickers
High Beta ($\beta \geq 1$)	NVDA, AAPL, AMZN, META, TSLA, AVGO, TSM, JPM, ORCL, MA
Medium Beta ($0 < \beta < 1$)	MSFT, GOOG, WMT, LLY, V, UNH, XOM, COST, PG, NVO
Negative Beta ($\beta \leq 0$)	ZM, BIIB, BEKE, ALAB, SMMT, ZTO, LINE, RBRK, MDGL, TAL

Figure 3 presents the scatter plots for the 2-week and 1-month testing experiments. Each data point in the scatter plots represents the RoR for an individual stock within the defined beta groups. The scatter plots highlight the variability in RoR across the beta categories for both testing durations. The 1-month experiment featured a larger dataset, reflecting the extended market exposure. Tables 10 and 11 summarize the detailed statistics for each beta group and the overall dataset for the 2-week and 1-month experiments, respectively.

Table 10. Detailed summary statistics and accuracy for overall and beta groups (2-week testing results).

Group	Mean	Median	Std Dev	Min	Max	95% CI	CV (%)	Skewness	Kurtosis	Accuracy (%)
Overall	4.17	2.90	4.38	−4.85	18.60	[2.77, 5.57]	105.00	1.02	1.57	91.67
High Beta	4.46	2.92	5.65	−4.85	18.60	[0.66, 8.25]	126.71	1.17	2.16	90.91
Medium Beta	3.14	2.79	2.27	−0.06	8.03	[1.77, 4.51]	72.28	0.82	−0.09	92.31
Negative Beta	5.02	4.08	4.93	−2.25	16.60	[1.89, 8.15]	98.12	0.90	0.66	91.67

Table 11. Detailed summary statistics and accuracy for overall and beta groups (1-month testing results).

Group	Mean	Median	Std Dev	Min	Max	95% CI	CV (%)	Skewness	Kurtosis	Accuracy (%)
Overall	3.61	3.09	4.08	−5.70	14.38	[2.53, 4.68]	113.02	0.95	1.05	89.74
High Beta	4.55	3.84	3.73	−4.23	14.38	[3.09, 6.01]	82.00	1.73	2.77	95.24
Medium Beta	1.92	1.70	2.39	−4.23	7.64	[0.70, 3.14]	124.48	−0.47	−0.72	83.33
Negative Beta	3.36	3.34	4.71	−5.70	14.28	[1.21, 5.51]	140.18	0.63	0.06	88.24

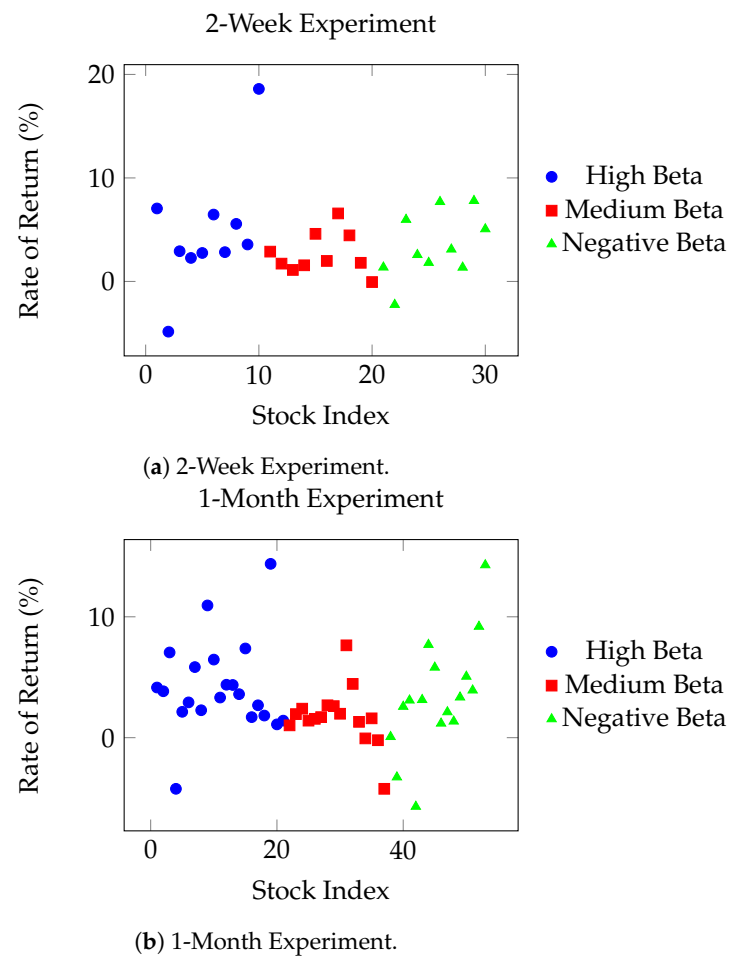


Figure 3. Side-by-side comparison of RoR scatter plots for 2-week and 1-month experiments.

For the 2-week experiment, the overall mean RoR was 4.17%, with the High Beta group achieving the highest mean RoR of 4.46%. Negative Beta stocks showed the greatest variability, with a standard deviation of 4.93%, but also recorded the highest accuracy at 91.67%. In the 1-month experiment, the overall mean RoR decreased to 3.61%, indicating a decline in returns over the extended period. The High Beta group maintained the highest mean RoR at 4.55% and achieved the highest accuracy of 95.24%, demonstrating consistent performance, despite the longer timeframe. The mean RoR for High Beta stocks improved slightly from 4.46% in the 2-week experiment to 4.55% in the 1-month experiment, while the Medium Beta group experienced the largest decline, with its mean RoR dropping from 3.14% to 1.92%. In terms of risk, Negative Beta stocks displayed the highest volatility, as indicated by their CV of 140.18% in the 1-month test. Despite this, High Beta stocks consistently delivered higher returns than Medium Beta stocks. Our system demonstrated robust predictive accuracy, achieving over 89% accuracy across all groups in both experiments. The highest accuracy was observed in the High Beta group during the 1-month experiment, underscoring the algorithm's effectiveness in predicting returns for volatile stocks.

Table 12 compares the performance of the beta groups against major market benchmarks. The benchmarks delivered significantly lower returns during the 1-month period compared to the beta groups. The High Beta group consistently outperformed all benchmarks, suggesting that high-beta stocks, while riskier, provided superior returns during the test period.

Table 12. Comparison of mean RoR for beta groups and benchmarks (2-week and 1-month testing results).

Category	Mean RoR (2-Week, %)	Mean RoR (1-Month, %)
Overall	4.17	3.61
High Beta	4.46	4.55
Medium Beta	3.14	1.92
Negative Beta	5.02	3.36
COMP	3.49	1.36
DJIA	4.79	0.40
S&P 500	3.97	1.10

The beta group categorization effectively highlights differences in performance, volatility, and prediction accuracy. High Beta stocks consistently outperformed benchmarks, making them an attractive option for investors seeking higher returns despite increased risk. These findings validated the robustness of our system across varying market conditions and time periods, emphasizing the value of combining technical analysis with sentiment-based predictions.

5. Conclusions

This study introduced a stock trading recommendation system that integrates sentiment analysis with technical indicators to generate actionable buy–sell recommendations. By combining seven days of financial news sentiment analysis with the RSI, the proposed system bridges the gap between traditional technical analysis and modern NLP techniques. Its computational efficiency and accessibility make it a highly practical tool for individual investors, operating effectively, without requiring high-end hardware or complex infrastructure.

The experimental results validated the system’s robustness and effectiveness across diverse market conditions. During the initial testing phase, the system achieved an average return of 4.03% across 95 stocks, representing a 5.10% increase in initial investment capital. The system consistently outperformed major market indices, including the COMP and S&P 500, demonstrating its ability to generate superior returns. In the market-cap-based classification experiment, the system demonstrated strong predictive performance, achieving an overall mean RoR of 4.76% (2-week) and 4.45% (1-month). Small Cap stocks performed best with 7.36% and 6.91%, while Large Cap stocks showed stability at 3.28% and 4.16%, and Mid Cap stocks underperformed in the 1-month period (2.27%). Accuracy was highest for Small Cap stocks (>90%), with Mid Caps showing greater variability. Compared to the benchmarks, the system’s 4.76% RoR (2-week) was slightly below DJIA (4.79%) but exceeded the COMP (3.49%) and S&P 500 (3.97%). Over 1 month, it significantly outperformed all indices, reinforcing its robustness and adaptability across stock categories. In a subsequent beta-based classification experiment, the DJIA index marginally outperformed the system in the 2-week period (4.79% vs. 4.17%), but High Beta stocks consistently led returns (4.46% and 4.55% over the 2-week and 1-month periods, respectively), surpassing all benchmarks. Negative Beta stocks showed high variability but strong returns, while the overall performance declined over the 1-month period (3.61%), reflecting short-term market momentum. Despite this, the system maintained over 89% accuracy, peaking at 95.24% for High Beta stocks, reinforcing its ability to capture high-return opportunities. Compared to the benchmarks, the system demonstrated superior predictive capability, validating its integration of sentiment-based and technical indicators.

This system offers significant value for individual investors by providing an automated and user-friendly solution for optimizing short-term trading decisions. By leveraging real-time sentiment analysis and technical indicators, it enables investors to make data-driven

decisions based on both quantitative and qualitative insights. The reliance on financial news data ensures objectivity, as news articles are typically curated and adhere to professional journalistic standards [25]. Unlike social media, which can be noisy and overly subjective, news data provide a consistent and reliable foundation for sentiment analysis, aligning with the system's design philosophy of minimizing bias.

The practical applications of this system are vast. Individual investors can implement it as part of their trading strategy by integrating it with widely available APIs from financial data providers and sentiment analysis tools. Its lightweight design ensures accessibility, even for investors without access to sophisticated trading platforms. Additionally, this system has the potential to be integrated into brokerage platforms, offering actionable insights directly within trading interfaces. For instance, personalized features such as customizable alert thresholds for sentiment shifts or RSI signals could further enhance its usability and alignment with investors' preferences.

Beyond individual investors, the system has broader implications for financial market analysis. The combination of sentiment analysis and technical indicators has been shown to provide complementary insights, improving the overall predictive power of trading systems [42]. By incorporating both short-term price patterns and real-time sentiment, the proposed system captures both market psychology and technical trends, making it especially effective in volatile or uncertain markets.

In conclusion, the proposed system addresses a critical demand for accessible and reliable trading tools that empower individual investors to navigate complex markets. Its lightweight design, high accuracy, and adaptability to diverse market conditions position it as a valuable tool for short-term trading strategies.

6. Future Work

While the proposed system demonstrates strong potential and practical value, it has certain limitations that provide opportunities for further research and development. Currently, the system is tailored for short-term investment strategies and does not explicitly incorporate long-term performance indicators such as earnings, revenue growth, or macroeconomic trends. Additionally, external factors like global events, policy changes, and unexpected market volatility are not fully addressed, which could impact the system's applicability to long-term investment scenarios.

Future iterations of the system will focus on integrating both short-term and long-term investment considerations. This includes incorporating fundamental financial metrics, macroeconomic data, and mechanisms to manage external influences effectively. To enhance its robustness, future developments will explore risk management features such as stop-loss orders and portfolio diversification strategies, broadening the system's applicability across various investment horizons.

Bias mitigation in sentiment analysis is another area for improvement. The current approach aggregates sentiment scores from up to 100 articles sourced from diverse outlets, reducing the impact of individual biases. However, eliminating biases entirely remains challenging. Future work will involve integrating source reliability metrics and weighting sentiment scores based on outlet credibility. Advanced techniques, such as de-biasing algorithms and cross-regional news sampling, will be explored to further improve the fairness and robustness of sentiment analysis.

In summary, future advancements will aim to enhance the system's versatility by addressing its current limitations and expanding its functionality. These improvements will make the system suitable for a broader range of investment strategies, empowering investors with a more comprehensive and reliable decision-making tool.

Author Contributions: Conceptualization, Y.K. and S.P.; Methodology, S.P.; Software, Y.K.; Writing—original draft, Y.K.; Writing—review and editing, S.Y. and S.P.; Supervision, S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The code supporting the findings of this study is publicly available on GitHub at the following repository (accessed on 9 January 2025): <https://github.com/KimYuri94/SimplifiedQuant/tree/main>. This repository includes the implementation of the proposed stock trading recommendation system, along with instructions for replicating the experiments described in this study. The financial data used in this study were obtained from Polygon.io and the sentiment analysis data were generated using NewsAPI. As these data sources are third-party services, users can access them by following their respective licensing and usage policies. No new datasets were created during this study. The research relied on publicly accessible financial and news data.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, F.; Yu, P.L.H.; Cheung, D.W. Complex stock trading strategy based on Particle Swarm Optimization. In Proceedings of the 2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr), New York, NY, USA, 29–30 March 2012; pp. 1–6. [\[CrossRef\]](#)
2. Gursida, H. The influence of fundamental and macroeconomic analysis on stock price. *J. Terap. Manaj. Dan Bisnis* **2017**, *3*, 222–234. [\[CrossRef\]](#)
3. Wang, H.; Wang, T.; Li, S.; Zheng, J.; Guan, S.; Chen, W. Adaptive long-short pattern transformer for stock investment selection. In Proceedings of the Thirty First International Joint Conference on Artificial, Vienna, Austria, 23–29 July 2022; pp. 3970–3977.
4. Ding, H.; Li, Y.; Wang, J.; Chen, H. Large language model agent in financial trading: A survey. *arXiv* **2024**, arXiv:2408.06361.
5. Vicari, M.; Gaspari, M. Analysis of news sentiments using natural language processing and deep learning. *AI Soc.* **2021**, *36*, 931–937. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Xing, F.Z.; Cambria, E.; Welsch, R.E. Natural language based financial forecasting: A survey. *Artif. Intell. Rev.* **2018**, *50*, 49–73. [\[CrossRef\]](#)
7. Jain, R.; Vanzara, R. Emerging Trends in AI-Based Stock Market Prediction: A Comprehensive and Systematic Review. *Eng. Proc.* **2023**, *56*, 254. [\[CrossRef\]](#)
8. Yu, S.J.; Yang, S.B.; Yoon, S.H. The Design of an Intelligent Lightweight Stock Trading System Using Deep Learning Models: Employing Technical Analysis Methods. *Systems* **2023**, *11*, 470. [\[CrossRef\]](#)
9. Huang, B.; Huan, Y.; Xu, L.D.; Zheng, L.; Zou, Z. Automated Trading Systems Statistical and Machine Learning Methods and Hardware Implementation: A Survey. *Enterp. Inf. Syst.* **2019**, *13*, 132–144. [\[CrossRef\]](#)
10. Aastha, S.; Arpit, J.; Prateek, S.; Sparsh, S.; Amrita, T. Sentiment Analysis of Stocks Based on News Headlines Using NLP. *AITEES* **2023**, *3*, 124–135. [\[CrossRef\]](#)
11. Zou, J. Machine Learning and Natural Language Processing in Stock Prediction. Doctoral Dissertation, School of Computer and Mathematical Sciences, University of Adelaide, Adelaide, Australia, 2023. Available online: <https://www.adelaide.edu.au/legals> (accessed on 9 January 2025).
12. Joshi, K.; Bharathi, N.; Rao, J. Stock Trend Prediction Using News Sentiment Analysis. *Int. J. Comput. Sci. Inf. Technol.* **2016**, *8*, 67–76. [\[CrossRef\]](#)
13. Yazdi, S.H.M.; Lashkari, Z.H. Technical analysis of Forex by MACD Indicator. *Int. J. Human. Manag. Sci. (IJHMS)* **2013**, *1*, 159–165.
14. Katayama, D.; Tsuda, K. A method of using news sentiment for stock investment strategy. *Procedia Comput. Sci.* **2020**, *176*, 1971–1980. [\[CrossRef\]](#)
15. Maqbool, M.; Shabir, M.; Iqbal, F. Stock prediction by integrating sentiment scores of financial news and MLP-Regressor: A machine learning approach. *Procedia Comput. Sci.* **2023**, *218*, 1067–1078. [\[CrossRef\]](#)
16. Lee, H.; Kim, J.H.; Jung, H.S. Deep-learning-based stock market prediction incorporating ESG sentiment and technical indicators. *Sci. Rep.* **2024**, *14*, 10262. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Fu, K.; Zhang, Y. Incorporating Multi-Source Market Sentiment and Price Data for Stock Price Prediction. *Mathematics* **2024**, *12*, 1572. [\[CrossRef\]](#)
18. Dahal, K.R.; Pokhrel, N.R.; Gaire, S.; Mahatara, S.; Joshi, R.P.; Gupta, A.; Banjade, H.R.; Joshi, J. A comparative study on effect of news sentiment on stock price prediction with deep learning architecture. *PLoS ONE* **2023**, *18*, e0284695. [\[CrossRef\]](#)
19. Shah, A. Stock Price Prediction Using Ensemble Model and Sentiment Analysis. *Int. J. Intell. Syst. Appl. Eng.* **2024**, *12*, 210. Available online: <https://www.ijisae.org/index.php/IJISAE/article/view/6720> (accessed on 9 January 2025). [\[CrossRef\]](#)

20. Investopedia. What Is a Stock Ticker Symbol? Available online: <https://www.investopedia.com/terms/s/stocksymbol.asp> (accessed on 5 October 2024).
21. Polygon.io. Available online: <https://polygon.io/> (accessed on 2 October 2024).
22. NewsAPI. Available online: <https://newsapi.org/> (accessed on 2 October 2024).
23. ISO 8601:2019; Date and Time Format. International Organization for Standardization: Geneva, Switzerland, 2019. Available online: <https://www.iso.org/standard/70907.html> (accessed on 9 January 2025).
24. Balahur, A. Sentiment Analysis in Social Media Texts. In Proceedings of the Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA), Atlanta, GA, USA, 14 June 2013; pp. 120–127. Available online: <https://aclanthology.org/W13-1617/> (accessed on 28 January 2025).
25. Schumaker, R.P.; Zhang, Y.; Huang, C.-N.; Chen, H. Evaluating Sentiment in Financial News Articles. *Decis. Support Syst.* **2011**, *49*, 308–318. Available online: <https://openknowledge.nau.edu/1478/> (accessed on 28 January 2025). [CrossRef]
26. Investopedia. What Is the Difference Between Fundamental and Technical Analysis? Available online: <https://www.investopedia.com/ask/answers/difference-between-fundamental-and-technical-analysis> (accessed on 26 January 2025).
27. Charles Schwab. How to Pick Stocks Using Fundamental and Technical Analysis. Available online: <https://www.schwab.com/learn/story/how-to-pick-stocks-using-fundamental-and-technical-analysis> (accessed on 26 January 2025).
28. Anderson, B.; Li, S. An investigation of the relative strength index. *Banks Bank Syst.* **2015**, *10*, 92–96.
29. OneUp Trader. Relative Strength Index (RSI) Indicator Explained. OneUp Trader. Available online: <https://blog.oneuptrader.com/trading-tips/relative-strength-index-rsi-indicator-explained/> (accessed on 26 January 2025).
30. Xiao, Q.; Ihnaini, B. Stock trend prediction using sentiment analysis. *PeerJ Comput. Sci.* **2023**, *9*, e1293. [CrossRef] [PubMed] [PubMed Central]
31. Loria, S. TextBlob: Simplified Text Processing; 2018. Available online: <https://textblob.readthedocs.io/en/dev/> (accessed on 2 October 2024).
32. Tasildar, R.B. Challenges in Using Abbreviations in Formal Writing and Newspaper Headlines. *Open J. Linguist. Commun.* **2024**, *2*, 45–55. Available online: https://www.wbnsou.ac.in/openjournals/Issue/2nd-Issue/July2024/5_RB_Tasildar.pdf (accessed on 28 January 2025).
33. Hamilton, W.P. *The Stock Market Barometer: A Study of Its Forecast Value Based on Charles H. Dow's Theory of the Price Movement*; Harper & Brothers: New York, NY, USA, 1922.
34. Granville, J.E. *Granville's New Key to Stock Market Profits*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1960.
35. TrendSpider. Official Daily Open and Close Prices: Their Importance and Impact on Stock Trading. TrendSpider Learning Center. Available online: <https://trendspider.com/learning-center/official-daily-open-and-daily-close-prices/> (accessed on 28 January 2025).
36. Cremers, M.; Petajisto, A.; Zitzewitz, E. Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation (January 21, 2010). EFA 2009 Bergen Meetings Paper, AFA 2010 Atlanta Meetings Paper. Available online: <https://ssrn.com/abstract=1108856> (accessed on 9 January 2025). [CrossRef]
37. Fama, E.F.; French, K.R. The cross-section of expected stock returns. *J. Financ.* **1992**, *47*, 427–465. [CrossRef]
38. Chan, L.K.C.; Chen, N.-F. Structural and return characteristics of small and large firms. *J. Financ.* **1991**, *46*, 1467–1484. [CrossRef]
39. Ritter, J.R. The long-run performance of initial public offerings. *J. Financ.* **1991**, *46*, 3–27. [CrossRef]
40. Investopedia. Beta Definition. Available online: <https://www.investopedia.com/terms/b/beta.asp> (accessed on 26 January 2025).
41. AvaTrade. Backtesting Trading Strategies. Available online: <https://www.avatrade.com/education/online-trading-strategies/backtesting-trading-strategies> (accessed on 26 January 2025).
42. Bollen, J.; Mao, H.; Zeng, X. Twitter Mood Predicts the Stock Market. *J. Comput. Sci.* **2011**, *2*, 1–8. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.