Original Research Article

Developing and testing a custom algorithmic trading strategy using exponential moving average, relative strength index, and sentiment analysis

Sstuti D. Mehra*, Sujala D. Shetty

Computer Science Department, Birla Institute of Technology and Science, Dubai Campus, Dubai 00000, United Arab Emirates

* Corresponding author: Sstuti D. Mehra, f20200029@bits-pilani.ac.in

ABSTRACT

Stock trading is a popular and important profession that requires near-to-perfect data analytical skills, mathematical and statistical knowledge, and a broad understanding of buying and selling stocks. Often, due to the number of factors to consider and the intervention of human bias, traders and investors make wrong decisions that cost them millions of dollars. Therefore, automated algorithmic trading has gained traction in the marketplace due to its ability to process huge amounts of data, perform mathematical calculations and make quick and effective decisions. Most algorithmic trading strategies rely on a single technical indicator; however, it has been found that combining two or more indicators makes a trading strategy profitable. Therefore, this paper proposes a custom algorithmic trading strategy that combines important technical indicators such as the Exponential Moving Average and Relative Strength Index and utilizes sentiment analysis of financial news as well. This combination of technical indicators and sentiment analysis is not prevalent in existing research. The performance of the strategy was tested on fifteen stocks from different sectors of the US market using Python’s VectorBt library. The results showed that most of the stocks produced a higher win rate with the custom strategy as compared to other strategies, with the highest win rate of 88% for the S&P 500 index. To carry out sentiment analysis, a NLP model using BERT was developed which achieved an accuracy of 84%. Finally, to test the strategy on real-time data, paper trading was carried out on the Alpaca API and after six months the portfolio’s ROI is 6.26%.

Keywords: stocks; algorithmic trading; trading strategies; exponential moving average; relative strength index; sentiment analysis; backtesting; paper trading

1. Introduction

According to reports from the International Monetary Fund (IMF) the global market was valued at $100 trillion by the end of 2022 indicating the exponential growth of the world economy. Trading in the stock market contributes not only to the global economy but to the wealth of corporations and individuals.

Stock trading involves individual investors or traders buying or selling shares of a public company, and it depends on multiple factors. While buying or selling stocks, individuals must carefully evaluate the market situation, observe the open, high, and close prices, and generate patterns to make an informed decision. This process can often be time-consuming, difficult, and affected by human emotions and bias.

Algorithmic trading involves the use of algorithms, specially
formulated trading strategies, and automated trading decisions to make stock trading efficient and reliable. Trading using a predefined strategy allows for a fast decision as computers can analyze huge amounts of data in seconds. This speed is crucial in an ever-changing market. Moreover, algorithmic trading eliminates human emotions and bias, which often negatively impacts traders. Trading using an algorithm can also help capitalize on market opportunities.

This paper studies technical indicators such as Relative Strength Index (RSI), Exponential Moving Average (EMA), and how these indicators display the stock market trend allowing traders to make decisions on buying and selling. A custom trading strategy is proposed which combines the results of both RSI and EMA on various stocks from different sectors of the US market and backtesting is performed to find the Win Rate for each strategy. To make the strategy more robust and reliable a Natural Language Processing (NLP) model has been developed to incorporate financial news sentiment analysis as Fazlija and Harder[1] found that sentiment scores are positively correlated to stock direction prediction. Finally, to perform trades using this custom strategy, paper trading on Alpaca API has been used to test the strategy.

The limitations of the existing research on algorithmic trading are the availability of effective trading strategies which incorporate a combination of important technical indicators and consider news sentiment analysis. Moreover, existing methods do not focus on backtesting the trading strategies, which is an important step to learn how the strategy would have performed in different market conditions in the past. Present research also does not conduct paper trading to test out their strategy’s performance in real-time market conditions.

In the paper, Section 2 explores the previous and related research work on algorithmic trading using technical indicators, the role of sentiment analysis and the limitations of existing studies. Section 3 describes the proposed methodology, discusses the working of EMA, RSI and backtesting the custom trading strategy on fifteen different stocks using VectorBt. This section also discusses the importance of including sentiment analysis and describes the working of the NLP model used. Section 4 includes the experimentation and results of backtesting the custom trading strategy and compares the win rates with other trading strategies using only single technical indicators. This section also discusses the results of conducting paper trading using the developed strategy on the Alpaca API. Section 5 includes conclusions and future work for the research.

2. Related work

With the advent of algorithmic trading, the formation of custom strategies that combine various technical indicators are gaining popularity. Technical indicators involving moving averages are generally based on the market momentum phenomenon which capitalizes a strategy’s returns in the past to predict future performance[2]. RSI works on the mean reversion theory. Mean reversion is a time series pattern which is described as the statistical tendency of an asset’s price to converge to the average price over time[3]. Traders are also considering news sentiment as stock prices are largely affected by company announcements and public opinion.

Authors Bajaj and Aghav[4] propose trading algorithms based on technical analysis. They make use of technical indicators such as RSI, EMA, and Average True Range to perform automated trading. With these indicators, the authors developed three combined trading strategies and tested them on the Nifty Index. These strategies outperformed trades using only one technical indicator. They achieved a win rate of up to 42% for their strategy using the three technical indicators together.

Ifleh and Kabbouri[5] design an approach which involves a correlation-based feature selection model to identify relevant technical indicators for different stocks. They made use of deep learning models such as ANN, CNN, and LSTM with a range of performance metrics. The authors found that ANN outperformed
other models for MASI, MADEX and FTSE 100 indices while CNN outperformed other methods for NASDAQ 100 and S&P 500.

Salkar et al.\cite{6} study that challenges in trading decisions can be significantly reduced by technical analysis of data. Technical indicators consider the price and volume of stock data and provide effective strategies for buying and selling stocks. The paper proposes trading strategies by combining various technical indicators based on quantitative analysis of time series data. The results show that a trading strategy with RSI and MACD technical indicators combined gives highest returns of 12%.

Sahin and Ozbayoglu\cite{7} study the performance of RSI for carrying out trades. They propose Trend-Normalized RSI which is a modification of the RSI by trend-removed stock data. The proposed model has several parameters like the trend detection period and buy/sell triggers. Genetic algorithms are used to optimize this model and the performance is compared against the general RSI. Results show that there is a performance improvement in profit and success rate using the new model. The authors conclude that the general RSI is extremely vulnerable to trend changes.

Seshu et al.\cite{8} conduct research that automated trading strategies based on different metrics can beat stock market baselines with higher returns. The authors propose a strategy that includes Bollinger Bands and Long Short-Term Memory (LSTM). The LSTM strategy uses predictions from 250 neural networks and the Bollinger Bands strategy uses close price, moving averages, and standard deviation to make buy/sell predictions. The performance is evaluated using historical data of the NIFTY50 index. Results show that custom strategies beat market baselines in 35.93% of all time periods tested and hence conclude that custom strategies produce higher returns.

Khedr et al.\cite{9} propose an effective model to predict stock market trends based on news sentiments and historical data analysis. The news sentiment is obtained using the Naïve Bayes algorithm and is combined with historical stock prices to make improved and accurate predictions. The proposed method achieves an accuracy of 89.8%.

Mehta et al.\cite{10} study the correlation of the movement of stock market prices with public opinions about a company. The authors proposed an algorithm that considered social media opinion, news, and historical stock prices and performed experiments using Support Vector Machine, Naïve Bayes classifier, and LSTM. The proposed methodology shows a positive correlation between sentiments and stock market movements. By using sentiment along with historical stock prices, a clear decision on buying/selling shares can be made.

Koroteev\cite{11} studies the applications of one of the most popular deep learning-based language models, Bidirectional Encoder Representations from Transformers (BERT). The paper describes the operations and applications of BERT. The author comes to a positive conclusion that BERT has represented a quantum leap in the field of NLP. This pre-trained model can effectively be used on large datasets for building intelligent algorithms.

3. Proposed methodology

The proposed custom strategy in Figure 1 includes using technical indicators such as Exponential Moving Average (EMA) and Relative Strength Index (RSI) on historical stock prices. This strategy is backtested using various stocks from different sectors of the US market. To make the strategy more dependable, a NLP model is developed which incorporates the BERT language model to classify daily news sentiments to finally give investors a signal on whether to buy or sell shares.
3.1. Technical indicators

Technical indicators are mathematical calculations performed on historical stock data to observe price trends and patterns. These trends and patterns can suggest favorable entry and exit points. Indicators can also identify overbought and oversold conditions and can confirm trading signals to assist traders in making informed decisions before buying or selling shares. According to research conducted by Faijareon and Sorniit[12], combining technical indicators produces more profitable trading strategies.

3.1.1. Exponential moving average

Exponential Moving Average (EMA) is like Simple Moving Average (SMA) (1). It measures the trend direction of the stock price and gives information on a good time to buy or sell shares.

EMA is stronger than SMA since the SMA takes an average of all the prices in each time period, however, EMA gives extra weightage to the recent stock prices, thereby giving a more accurate trend (Funde, Damani[13]). The EMA is more responsive to latest price changes than the SMA, which usually shows a lag in the trend.

\[
SMA = \frac{A_1 + A_2 + \ldots + A_n}{n}
\]

\[A_n = \text{price of an asset during time period n}\]

\[n = \text{total number of periods}\]

\[
EMA_{today} = \left( Value_{today} \times \left( \frac{\text{Smoothing}}{1+\text{Days}} \right) \right) + EMA_{yesterday} \times \left( 1 - \left( \frac{\text{Smoothing}}{1+\text{Days}} \right) \right)
\]

In Equation (2), smoothing is usually assigned a value of 2. Smoothing gives the most recent observations the most weight and as it is increased more recent observations have more influence on the EMA.

For the custom strategy, a fifty-day EMA and a two-hundred-day EMA are used to identify uptrends and downtrends. If EMA50 crosses above EMA200, it indicates a bullish or uptrend and if EMA50 crosses below EMA200, it indicates a bearish or downtrend.
In Figure 2, we can observe that when EMA50 crosses over EMA200 (August 2019, May 2020) the stock price starts showing a rising trend and when EMA50 crosses below EMA200 (November 2018, March 2020) the price starts showing a falling trend.

3.1.2. Relative strength index

The Relative Strength Index or RSI is a popular technical indicator used for effectively trading shares in the stock market. RSI is a momentum indicator based on mathematical calculations, which can help trading analysts observe if the market is bearish (overbought) or bullish (oversold).

\[
RS = \frac{\text{Average Gain}}{\text{Average Loss}}
\]

(3)

\[
\text{RSI} = 100 - \frac{100}{1+RS}
\]

(4)

In Equation (3), the Relative Strength (RS) is calculated by obtaining the average gain and loss in the stock prices, which is the difference between the daily open and close prices and in Equation (4) the index is calculated.

Most commonly the RSI is calculated for fourteen days period and is mapped on a scale of 0 to 100. As shown in Figure 3. If the RSI value crosses below 30, it means the market is in an oversold condition and it is a good time to buy shares, whereas, if the RSI crosses above 70, it indicates that the market is in an overbought condition, and it is a good time to sell the shares\textsuperscript{[14]}.

3.2. Backtesting using vectorbt

Backtesting is a process to evaluate the performance of a certain trading strategy in different market conditions. It involves applying the developed trading strategy on historical data to simulate how the strategy
would have performed. It follows the assumption that if a strategy worked well in the past, it has a good chance of working well for future circumstances and vice versa\textsuperscript{[15]}.

With trading strategies becoming more complex and involving analysis of large amounts of data, automated backtesting is crucial. Backtesting is an important step for brokers and investors to assess if trading strategies are profitable in the marketplace and for different kinds of stocks. It helps in performance comparisons with other strategies and helps investors set realistic expectations, by analyzing the strategies’ potential returns and win rate.

To perform comprehensive backtesting on the proposed custom strategy using EMA and RSI, a Python library called VectorBt has been used. VectorBt is a Python package that allows for the testing of multiple trading strategies in seconds. It performs superfast computations using vectorized operations with NumPy and non-vectorized operations with Numba.

To conduct extensive backtesting on the custom strategy, important stocks from the eleven sectors of the US stock market, as defined by the Global Industry Classification Standard (GICS) have been considered\textsuperscript{[16]}. The eleven GICS stock market sectors include:

- Energy
- Materials
- Industrials
- Utilities
- Healthcare
- Financials
- Consumer Discretionary
- Consumer Staples
- Information Technology
- Communication Services
- Real Estate

A time period of ten years from 2012 to 2022 has been considered while backtesting. This enables a complete performance evaluation of the custom trading strategy.

3.3. Sentiment analysis using bert

According to the study by author Wojarnik\textsuperscript{[17]}, analyzing the sentiments of individuals from a discussion forum proved useful in determining when to buy or sell shares for both short-term and long-term investments. Human sentiments affect human behavior and thus stock markets are influenced by its effects\textsuperscript{[18]}.

Sentiment analysis is an important factor to consider for trading strategies. Stock markets are vulnerable to latest news developments, public opinion, and the reputation of corporations. Therefore, to develop a more robust strategy, daily news sentiments have been analyzed using deep learning to help investors make informed decisions before trading. If the news sentiment for a stock is positive it means the company is performing well and its stock price is going to rise, then it is a favorable time to buy. If the sentiment is negative then investors get an idea that the stock is not performing well or its price is going to drop, which makes it a good time to sell shares and make a profit.

To obtain the most accurate sentiment predictions, a pre-trained model called BERT (Bidirectional Encoder Representations from Transformers) has been utilized.

BERT is an open-source deep-learning language model in which every output element is connected to every input element, and the weights in between them are dynamically calculated. BERT is a Masked Language Model (MLM) which enables bidirectional learning from the input text. Therefore, it is a pre-trained bidirectional model which predicts a masked word by considering both the next token and the
previous token of the word. The model also consists of the Transformer architecture which makes BERT feasible to train on large amounts of text data\cite{19}.

As depicted in Figure 4, for this study a NLP model is built using three Input layers in the Keras layer which is passed into a Dense Layer with a Sigmoid activation function to make a classification prediction. The BERT layer is included along with the Keras layer and provides the tokenized and masked inputs to the model.

![Figure 4. NLP model architecture.](image)

### 4. Experimentation and results

#### 4.1. Datasets

To calculate the EMA and RSI and backtest the strategy, the historical stock prices for all stocks have been obtained from Yahoo Finance. The datasets contain the date, open and close prices, high and low prices, and the trade volume for a particular day.

Training and validation of the NLP model using BERT has been done using the Sentiment Analysis for Financial News dataset by authors Malo et al.\cite{20}. The dataset contains over 4800 rows of news headlines and their corresponding sentiment of either positive, neutral, or negative.

To carry out paper trading on the final strategy using the Alpaca API, real-time stock prices and financial news has been obtained from the Polygon.io API.

#### 4.2. Results

##### 4.2.1. Backtesting

Extensive backtesting has been carried out on eleven stocks from different sectors of the US stock market to understand and evaluate the performance of the custom strategy in various marketplaces. S&P 500 and NASDAQ 100 indexes have also been included since they cover a wide variety of stocks and set a benchmark for the US stock market. Many investors use these indices to evaluate the performance of their portfolios.

To measure the performance of the strategy, the win rate has been examined. Win rate Equation (5), also known as success rate, is the percentage of winning trades out of the total number of trades executed in a given time period.
From Table 1 we can observe that for most sectors such as materials, industrials, healthcare, consumer discretionary, consumer supplies and communication services, and for the indices of S&P 500 and NASDAQ 100, the custom strategy of EMA and RSI has a higher win rate than the strategies which use only EMA or only RSI.

![Image](image_url)

\[
\text{Win Rate} = \frac{\text{Number of winning trades}}{\text{Total number of trades}} \times 100
\]

This shows that for many sectors of the US stock market, a combination of EMA and RSI produces better decisions for buying and selling shares. The S&P 500 and NASDAQ 100 have the highest win rates of 88% and 77.2% respectively, which means that the custom strategy will almost always prove most favorable for indices that represent a diversified group of stocks.

It can also be observed that using only EMA for making trading decisions has the lowest win rate for all fifteen stocks studied. This highlights the weaknesses of using EMA alone and further proves the point of the effectiveness of combining technical indicators.

While using only RSI does produce higher win rates for some sectors of the stock market, it can often be risky to rely on just one technical indicator to make buying and selling decisions in the stock market.

The backtesting analysis for Alphabet stocks from the years 2012 to 2022 has been shown in Figure 5. Considering the market conditions using EMA50, EMA200 and RSI calculation, the green upward arrows indicate the execution of buying orders and the red downward arrow indicates when the shares would have been sold. The win rate for this combined strategy is 66.7% which is higher than the win rates of 59.1% and 62.5% when using only EMA and only RSI in trading strategies respectively.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Stock Name</th>
<th>Stock Ticker</th>
<th>Win Rate (%)</th>
<th>EMA</th>
<th>RSI</th>
<th>EMA &amp; RSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>ExxonMobil</td>
<td>XOM</td>
<td>0</td>
<td>44</td>
<td>41.3</td>
<td></td>
</tr>
<tr>
<td>Materials</td>
<td>DuPont</td>
<td>DD</td>
<td>60</td>
<td>63.8</td>
<td>67.7</td>
<td></td>
</tr>
<tr>
<td>Industrials</td>
<td>Raytheon</td>
<td>RTX</td>
<td>40</td>
<td>48.1</td>
<td>48.2</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>Duke Energy</td>
<td>DUK</td>
<td>50</td>
<td>70</td>
<td>63.6</td>
<td></td>
</tr>
<tr>
<td>Healthcare</td>
<td>Pfizer</td>
<td>PFE</td>
<td>42.8</td>
<td>44.4</td>
<td>63.6</td>
<td></td>
</tr>
<tr>
<td>Financials</td>
<td>JP Morgan Chase</td>
<td>JPM</td>
<td>57</td>
<td>73</td>
<td>66.7</td>
<td></td>
</tr>
<tr>
<td>Consumer Discretionary</td>
<td>Amazon</td>
<td>AMZN</td>
<td>66.6</td>
<td>75.1</td>
<td>77.7</td>
<td></td>
</tr>
<tr>
<td>Consumer Staples</td>
<td>Proctor &amp; Gamble</td>
<td>PG</td>
<td>30</td>
<td>53.8</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Information Technology</td>
<td>Apple</td>
<td>AAPL</td>
<td>60</td>
<td>66.6</td>
<td>64.5</td>
<td></td>
</tr>
<tr>
<td>Communication Services</td>
<td>Alphabet</td>
<td>GOOGL</td>
<td>59.1</td>
<td>62.5</td>
<td>66.7</td>
<td></td>
</tr>
<tr>
<td>Real Estate</td>
<td>Simon Property Group</td>
<td>SPG</td>
<td>40</td>
<td>57.1</td>
<td>56.6</td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>S&amp;P 500</td>
<td>SPY</td>
<td>60</td>
<td>82.1</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>NASDAQ 100</td>
<td>QQQ</td>
<td>66.6</td>
<td>70.3</td>
<td>77.2</td>
<td></td>
</tr>
</tbody>
</table>
4.2.2. Sentiment analysis

The developed NLP model using BERT was trained and validated with an eighty-twenty split of the dataset. Three epochs and a batch size of 32 were used. The final accuracy after training and validating the model three times came to 84%. This shows that the NLP model using BERT can almost accurately predict news sentiments.

4.3. The alpaca API

To test out the custom strategy and check if it responds to the market conditions correctly, the Alpaca API[21] has been used.

The Alpaca API offers a commission-free paper trading platform, where users can test their strategy on real-time stock market data. Using the API, users can readily submit orders to buy or sell stocks and view their open positions. The platform offers crypto, broker, and stock trading.

The submitted orders are reflected on the account’s dashboard where users can see their equity, buying power, and cash available. Initially, a paper trading account contains $100,000 cash. Further on the dashboard, users can view their recent orders and how their portfolio is performing in the stock market.

To evaluate the custom strategy using EMA, RSI, and sentiment analysis, the latest stock price and news data were obtained for Polygon.io API for Alphabet stocks.

Using Python, EMA, RSI were, and the news sentiments were obtained for Alphabet stocks. When the fast EMA (EMA50) crossed above the slow EMA (EMA200) and when the RSI value crossed below 30 and the news sentiment was positive, a buying order for 100 shares was submitted.

Similarly, a condition for a selling order was executed when EMA50 crossed below EMA200 and the RSI value crossed above 70 and the news sentiment was negative.

When neither of these conditions were met, no order was executed. These changes were reflected on the Alpaca dashboard and provided valuable details of the price at which the shares were bought and sold.

Figure 5. Indication of buy and sell trades for alphabet stocks over a period of ten years after conducting backtesting of the custom strategy.
Over the course of six months, the portfolio executed buy orders twice and a sell order once. To calculate the returns of the custom strategy, the Return on Investment (ROI) metric has been utilized. ROI Equation (6) is a ratio which divides net profit or loss on an investment by the initial cost of the investment. The Alphabet stocks were initially bought for a price of $12,380 and the price after six months stands at $13,156. Therefore, the ROI from the custom algorithmic trading strategy is 6.26% in a timeframe of just six months which further proves the success of the strategy.

\[
\text{ROI} = \frac{\text{Net Income}}{\text{Initial Cost of Investment}} \times 100 \quad (6)
\]

5. Conclusions and future work

In this study, a custom trading strategy using EMA, RSI and news sentiment analysis was proposed. To evaluate the performance and effectiveness of the strategy, it was tested on eleven stocks from different sectors of the US market as well as the S&P 500 and NASDAQ 100 indices. Results show that six out of the eleven stocks have a higher win rate than other trading strategies using only EMA or only RSI. The custom strategy produced the highest win rates for both indices which proves that the proposed strategy works well with a wide variety of stocks.

The custom strategy also included developing a deep learning NLP model using the BERT language model to perform accurate sentiment analysis on financial news. The model obtained an accuracy of 84%. Finally, this strategy was tested on the Alpaca API, a paper trading platform. The portfolio executed automated buying and selling orders using the custom strategy and the value of shares are profitable after two months of testing.

Future work for this study includes training a machine learning model to predict when to buy and sell shares based on the performance of the proposed custom strategy, to further strengthen the reliability of the strategy and make it accessible to brokers and investors.

Author contributions

Conceptualization, SDM, SDS; methodology, SDM; software, SDM; validation, SDM, SDS; formal analysis, SDM; investigation, SDM; resources, SDM, SDS; data curation, SDM; writing—original draft preparation, SDM; writing—review and editing, SDS; visualization, SDM; supervision, SDS; project administration, SDS; funding acquisition, SDS. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

I would like to express my gratitude towards Birla Institute of Technology and Science, Dubai Campus for providing me with the opportunity and resources such as access to research papers and material from the college library to carry out an extensive study.

Conflict of interest

The authors declare no conflict of interest.

References


5. Ifleh A, El Kabbouri M. Stock price indices prediction combining deep learning algorithms and selected technical indicators based on correlation. Arab Gulf Journal of Scientific Research. Published online October 17, 2023. doi: 10.1108/agjsr-02-2023-0070


