Developing non-Machine Learning Algorithms for determining when to buy and sell on the stock market

Overview

This project is centered around developing algorithms which could be used to help individual investors in the US participate more effectively in the stock market. This is especially crucial due to the fact that around 60% of Americans invest in the stock and it holds over 40% of US wealth. Recently, large firms have been able to significantly outperform the market, meaning that smaller investors have underperformed and thus have made the stock market a far worse option for investment.

Significance

The stock market is a place where fortunes are won or lost depending on how accurate the predictions of price movement are. A very important part of these predictions is determining when to buy and sell stocks. This is often done through algorithms, as these can counteract the natural human desire to develop storylines. For example, a study found that investors “were influenced by properties of both news items and price series but they relied more on the former” (Sobolev et al., 2017). This may be due to the fact that most investors don’t have access to algorithms to help them make sense of the data when investing

Table 1. Results of Experiment 1

			Trend	
			Positive	Negative
Panel A : Western group, N = 30				
Trading latency	News valence	Good	48.33 (48.89)	60.00 (61.84)
		Bad	45.67 (38.28)	36.67 (24.75)
Share number	News valence	Good	1.35 (0.92)	0.83 (0.96)
		Bad	1.03 (1.01)	0.4 (0.81)
Returns	News valence	Good	3.14 (2.00)	−3.31 (2.22)
		Bad	2.12 (1.38)	−2.68 (1.16)
Panel B: Eastern group, N = 30				
Trading latency	News valence	Good	92.33 (71.98)	106.34 (70.78)
		Bad	66.00 (55.27)	73.66 (61.45)
Share number	News valence	Good	1.13 (0.96)	1.15 (0.917)
		Bad	0.72 (0.96)	0.60 (0.87)
Returns	News valence	Good	4.24 (2.25)	−4.35 (2.25)
		Bad	2.94 (1.89)	−3.44 (2.02)

Table 1 (Taken from (Sobolev et al., 2017): This table shows the reaction to news and data, demonstrating the overreliance on news

Although taking current events and news into account is an important component of stock prediction, looking at historical trends tends to provide a better mathematical base to investments. Most algorithms currently in use today are complex and utilize complicated machine learning methods, meaning that they are developed and kept secret by very large investors. This restricts these algorithms to only the uber-rich and investment firms, giving them a leg up over smaller investors for whom these machine learning algorithms are too cumbersome to utilize effectively. To this end, small investors are mostly left to use very basic strategies such as buy and hold, leaving them at a significant disadvantage. Developing an algorithm that is more complicated than buy and hold but simpler than the machine learning algorithms used by large firms will help fill a noticeable gap in the industry.

Background Information

In this paper, the focus will be on three different investment tools: stocks, mutual funds and exchange-traded funds. A stock is a fractional part of a company, and the owner of a stock is a shareholder in that company, meaning that they own some portion of that company. A mutual fund is an actively managed “investment vehicle

that pools money from multiple investors to purchase a diversified portfolio of stocks, bonds, or other securities (according to the fund's stated strategy)" (Picardo, 2024). Finally, exchange-traded funds (ETFs), are a hybrid of stocks and mutual funds, where they can be traded at any time like a stock and usually track a single index, while still owning small parts of a number of different stocks.

The strategy that is usually recommended for small investors is buying and holding a single stock or fund for an extended period, as this generally flattens out market fluctuations over time and usually leads to growth in the long term. However, the issue with these strategies is that the maximum return on investment is quite low, and it also may prevent the investor from selling stocks that are losing value. This strategy is very far from the much more robust and active management strategies used by larger investors, as these strategies tend to be somewhat riskier but usually significantly more effective. These strategies aren't even viable for smaller investors that will tolerate risk as they will need to spend large amounts of time "analyzing financial statements" and "to keep up to date with recent news about industries" (Taylor, 2024).

Previous Research

Most previous research conducted in this field has focused on machine learning algorithms that used the previous trends of the specific investment to attempt to predict its future movement. There has been relatively little research focused on making algorithms without machine learning due to the assumption that a machine learning algorithm will be more accurate. It was also assumed that the buy and hold approach should be sufficient for those smaller investors whom the machine learning algorithms didn't suit. However, recent research has shown that the buy and hold strategy may not be as viable in the new era of stock market volatility (Sanderson & Lumpkin-Sowers, 2018). The worsening outcomes of the buy and hold approach suggest that a new strategy will be necessary for low-volume investment.

Table 2. Success Rates for All NASDAQ Exchange Traded Funds (ETF's).

Holding Period (Years)	Mean Prob. of a Gain	Min	Max	Range	Std Dev	Skew	CV	Number of ETF's	Survival Rate
1	59.49%	1.30%	100.00%	98.70%	24.03%	−0.5792	0.404	1374	
2	67.12%	1.16%	100.00%	98.84%	27.43%	−0.6458	0.409	1228	89.37%
3	72.71%	1.22%	100.00%	98.78%	26.86%	−0.8895	0.369	1055	76.78%
4	75.18%	1.18%	100.00%	98.82%	26.85%	−0.9540	0.357	944	68.70%
5	78.29%	2.74%	100.00%	97.26%	25.83%	−1.1139	0.330	799	58.15%
6	83.73%	4.55%	100.00%	95.46%	24.76%	−1.5002	0.296	673	48.98%
7	87.11%	6.67%	100.00%	93.33%	22.88%	−1.8630	0.263	571	41.56%
8	87.24%	3.19%	100.00%	96.81%	23.60%	−2.0863	0.271	452	32.90%
9	93.20%	9.00%	100.00%	91.00%	16.12%	−3.0983	0.173	332	24.16%
10	95.97%	10.00%	100.00%	90.00%	11.97%	−5.0068	0.125	207	15.07%
11	97.02%	40.79%	100.00%	59.21%	8.81%	−4.0153	0.091	166	12.08%
12	96.90%	43.75%	100.00%	56.25%	8.80%	−3.7230	0.091	138	10.04%
13	97.71%	55.77%	100.00%	44.23%	7.71%	−4.0150	0.079	112	8.15%
14	98.01%	50.00%	100.00%	50.00%	7.59%	−4.4286	0.077	101	7.35%
15	97.38%	36.36%	100.00%	63.64%	10.29%	−4.2739	0.106	82	5.97%
16	97.98%	66.67%	100.00%	33.33%	7.28%	−3.8001	0.074	32	2.33%
17	99.17%	83.33%	100.00%	16.67%	3.35%	−4.2810	0.034	30	2.18%
18	98.96%	79.17%	100.00%	20.83%	4.66%	−4.4721	0.047	20	1.46%
19	99.12%	83.33%	100.00%	16.67%	3.82%	−4.3589	0.039	19	1.38%
20	100.00%	100.00%	100.00%					2	0.15%
21	100.00%	100.00%	100.00%					1	0.07%
22	100.00%	100.00%	100.00%					1	0.07%
23	100.00%	100.00%	100.00%					1	0.07%

Table 2 (Taken from Sanderson & Lumpkin-Sowers, 2018): This is a demonstration of the results of a 19 year buy and hold strategy applied to ETFs, and it shows that only .07 % survive due to the quick turnover, and thus a simple buy and hold is almost 100% losing over long periods of time

Problem Statement

Objective

This project' objective is to develop a non-machine learning algorithm to determine when to buy and sell on the stock market and compare it to previous such algorithms and see improved performance. Our long-term goal is to develop an algorithm that will outperform previous non machine learning

algorithms to be usable for small investors. This will slightly level the playing field for small investors in comparison to larger firms and make the stock market a place where investors have similar chances of profitability or loss no matter their size.

Specific Aim 1: The first aim of the project was to develop a number of algorithms for determining when to buy and sell using a variety of different techniques.

Specific Aim 2: The second aim of the project was to determine how well these algorithms performed against conventional algorithms and compare them to each other in order to find the strongest algorithm to hopefully compare it to ML algorithms.

The expected outcome of this work is to create a non-ML algorithm that is relatively simple to understand for those with a high-school level understanding of mathematics and a solid understanding of the stock market in general. Additionally, this algorithm is ideally meant to outperform the buy and hold algorithm, and be somewhat comparable to machine learning algorithms, which make correct decisions roughly 70% of the time (Ayyildiz & Iskenderoglu, 2024).