Original article

Journal of High School Science

Predictive analysis of the S&P 500 stock price index using Machine Learning

Lee S[1], Sung Z[2]

Submitted: March 13, 2024, Revised: version 1, May 22, 2024, version 2, May 22, 2024, version 3, June 4, 2024
Accepted: June 16, 2024

Abstract
We attempted to predict the future trend of the S&P 500 (Standard and Poor's 500) stock market index using machine learning algorithms. Unlike the fundamental analysis which is affected by intangible economic circumstances, we hypothesized that using only technical analysis, we could effectively predict future index movement based on historical trends if the targets were well defined within a period of the movement. This technical approach takes advantage of the indicators defined by mathematical or statistical measures, called the features. We chose price, volume change, day gain, 50- and 200-day Simple Moving Average (SMA), Relative Strength Index (RSI), and Moving Average Convergence Divergence (MACD), as the features. Initially, the S&P 500 index between Jan 19, 1993, and Feb 17, 2023 was represented with these features. After splitting the index into training and learning sets, the future trends were predicted with two supervised classification algorithms; decision tree and random forest, for the targets. We set the targets as next day, next week, and next month. We found that prediction accuracy increased as the defined prediction period increased. For example, the prediction accuracy for the next month was ~ 80%, as compared with ~ 50% for that of the next day. Prioritizing the indicators' importance suggested that the SMA contributed the most to prediction accuracy. In addition, we investigated serval stock price trends of companies representing sustainable environment engineering, finance, and Information Technology in the commercial market using the same approach. These individual stocks trended comparable with the S&P 500 stock index analysis. Since SMA is a process to smooth out the trends over long transitions, it is less sensitive to time-related variation and capable of predicting long-term trends with minimal indefinite fluctuations. Thus, as the prediction horizon increased, the indicators that projected longer trends became more critical.

Keywords
Machine Learning, Data science, Technical analysis, S&P 500 index, Simple Moving Average, Relative Strength Index, Moving Average Convergence Divergence, Market prediction, Decision tree, Random forest

[1]Sihyun Lee, Asia Pacific International School (APIS), 57 Wolgye-ro 45ga-gil, Nowon-gu, Seoul, 01874, South Korea. sihyunlee1010@gmail.com
[2]Corresponding author: Zuhawn Sung, Fermilab, Batavia, IL 60510, USA. zsung@fnal.gov

1. Introduction

Investing in stocks is a convincing way of building wealth, as it allows one to partake in the financial gains of growing companies (1). High returns and high rewards in the stock market are a strong desire for investors. Investing is a substantial part of economic activities of capitalism (1). The total global capitalization of the stock market for all publicly traded companies in 2020 was valued at approximately US $93 trillion (2). There are various options investing in the stock market and information and guidelines are readily available. Anyone can design any investment style based on individual preferences and knowledge, to suit their own purpose. However, traditional prediction is often unproductive due to the large volume of transactions and information. Thus, significant time and effort are required for high returns. Trading practices have evolved to implementing algorithms to deal with large amounts of information, targeting optimal productive investment. Among these are the so called "AI (artificial intelligence)" algorithms that can imitate human brain capability, likely learning the previous data and predicting the unknown results from training. Machine learning (ML) is the processing part of AI. ML makes it easier for investors to create models that take into account the complexity of the stock market environment, such as GDP growth, inflation, depression, etc. (3, 4).

Among one of the more popular stock market index that investors follow in the US is the S&P 500 (Standard and Poor's 500). This index tracks the stock performance of the top 500 selected companies listed on the stock exchange (5). The index is a joint venture majority-owned by S&P Global. Conference Board Leading Economic Index (an American economic leading indicator) applies the market index as a computing factor. It describes significant points of the movement during a business cycle. S&P Dow Jones indices, simply the Dow, maintains the S&P 500 index. The index's average annualized return has averaged 11.8% since its inception in 1928 through December 31, 2021 (6).

There are two representative styles of stock investments (7). First, the investors with a passive strategy intend to benefit from long-term market price increases. This is a buy-and-hold strategy that minimizes investors' mistakes from reacting emotionally to every stock market move. The second style of investing is the active strategy. This style places significant emphasis on the expertise of experienced stock analysts, so that information acquired by the analysis is used to buy and sell stocks with specific characteristics (8). Stock market analysis is divided into two different disciplines: "fundamental analysis" and "technical analysis" (9). The fundamental analysis predicts a company's future stock prices based on the business environment and financial performance. It is a method to forecast time-sensitive stock values. Its objective is to fit into classical problem-solving in econometrics. Intangible factors like the economy, politics, interest rates, and mass media mainly contribute to this analysis with the company's stock market and business. The fundamental analysis assesses potential investments in tandem with economic and trend indicators. It represents the economy's intrinsic characteristics and considered beneficial for long-term investment. On the other hand, technical analysis predicts the stock market based on an evaluation of charts and statistical graphs associated with previous market trends. The expected future market behavior can be predicted related to the past performance in the price changes and profits. The technical analysis is a quick assessment and fast decision-making process. It only takes into account, statistical or mathematical measures (10).

Machine Learning (ML) evaluates large datasets effectively and then provides accurate forecasts from learning data *via* optimized models. ML finds variables for accurate prediction (11). In financial analysis, the models, called algorithms, provide spotted patterns over the entire trade transaction. The ML models allow the investors to make informed judgments for high returns. In this study, we predicted the future trend of the S&P 500 stock index (12) in terms of the technical analysis using ML algorithms. We hypothesized that the technical analysis could provide time-insensitive accurate predictions compared to the fundamental analysis, leading to steady prediction over long-period transactions. This assumption differs from the general practice of the fundamental analysis. This analytical approach does not reflect frequent see-sawing trends reflected by economic environments that emerge quickly. We used a series of periods such as day-to-day, week-to-week, and month-to-month as the target. After the S&P 500 index analysis, we also investigated the suitability of our technical approach to the stock price market for three RE100 companies.

2. Materials and Methods

2.1. *Indexes and features*

The historical data of the S&P 500 ETF Trust (SPY, the NYSE ticker symbol) was used in this study. This dataset was downloaded *via* the Yahoo finance website (18). ETF stands for Exchange Traded Funds, a fund tracking asset. This fund holds various underlying assets, such as a particular index, sector, or commodity. Historical stock data provides information on price movements, range, volume, and other stocks. In contrast, the daily price represents the OHLCV (open, high, low, close, & volume) reported by the common stock's principal exchange or quotation system. The attributes of the historical data are open, high, low, and close (OHLC), adjusted close, and volume. There are various methods to calculate indicators directly or indirectly. Figure 1 shows an early time regime of OHLCV of SPY, and Figure 2 presents the price trends in dollars since 1993-1-29.

The next day, next week, and next month were chosen as the targets to evaluate the future trend of the S&P 500 using technical analysis. These times represent the progressive extension of a target horizon. We applied various stock trending indicators for prediction modeling, implemented as features in machining learning algorithms. These indicators, as described earlier, were price change, volume change, day gain, 50-days simple moving average, 200-days SMA, RSI, MACD, and MACD signal (19). These are heuristic pattern-based signals used for decision-making and useful to identify trends but cannot independently predict trends (20, 21).

Date	Open	High	Low	Close	Adj Close	Volume
1993-01-29	43.96875	43.96875	43.75000	43.93750	25.218216	1003200
1993-02-01	43.96875	44.25000	43.96875	44.25000	25.397573	480500
1993-02-02	44.21875	44.37500	44.12500	44.34375	25.451387	201300
1993-02-03	44.40625	44.84375	44.37500	44.81250	25.720436	529400
1993-02-04	44.96875	45.09375	44.46875	45.00000	25.828049	531500

Figure 1. A part of the SPY dataset from 1993-01-29. This dataset is downloaded from the Yahoo website for a daily price.

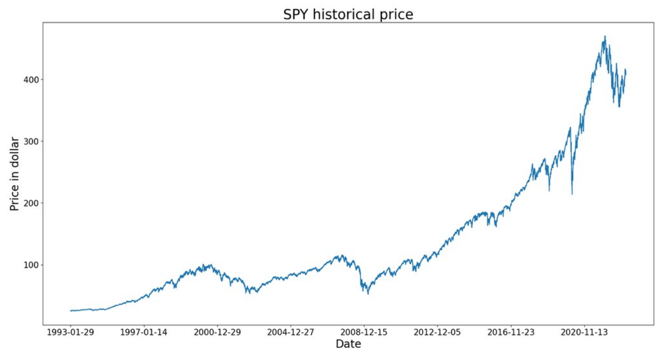

Figure 2. The SPY historical price from 1993-01-29 to 2023-02-17. This plot shows the adjusted close price in dollars.

We used information up to a particular day to predict the price for the next day, next week, and next month. The information for a particular day could be collected for up to 200 days, for example, to calculate the 200 day moving average of price. The features we used for a particular day were calculated based on that day's information or the past relevant time frame; such as price changes between today and yesterday, volume changes between today and yesterday, 50 day moving average, 200 day moving average, etc. How far back we went in time thus depended on the availability of the 200 day moving average, which was the longest span in our features. We started using data from 2011 and filtered out the initial 200 days since the 200 day moving average was not available during those initial times. For our second prediction point, we used the previous 31 days to predict the price at time 32 days, time 38 days and time 61 days. In other words, we used sliding windows. The rationale behind this was to make the method executable on any day when features are available. This would make the algorithm useful whenever there is opportunity for positive gain for the next day, week or month (depending on the trading period).

Price and volume change describe the difference in the closing price between today and yesterday. The Day gain represents the variation between opening and closing during a day. The SMA is the sum of all the closing prices of a security over a certain number of periods, which is then divided by the total number of periods. The SMA shows smoothed-out fluctuations of the daily price with a constant updated average. SMA is a simple but powerful visualization tool to trace spotted patterns. If SMA shows signals of increasing or decreasing, the stock is in an uptrend or *vice versa*. SMA of 50- and 200-days are used as arithmetic means of a given set of prices over the specified days. SMA can be defined as the following formula (22).

$$SMA = \left[A \middle| 1 + A_2 + A_3 + \ldots. + A_n \right] / n$$

A is the stock price in period n, and n is the number of time periods. On the other hand, an exponential moving average (EMA) is applied for MACD signals, as described (23).

$$EMA_t = \left[V_t \times \left(\frac{s}{1+d} \right) \right] + EMA_y \times \left[1 - \left(\frac{s}{1+d} \right) \right]$$

Where EMA_t and EMA_y are the EMA of today and yesterday, respectively. s is a smoothing factor, and d is the number of days. This EMA has the ability to rapidly discern changes in responses and has a higher value when the stock price rises and falls faster when the price declines.

The moving average convergence divergence (MACD) is an indicator that follows trend momentum to compare two moving averages and dissect their relationship. The MACD is described by subtracting the 26-period exponential moving average (EMA) from the 12-period EMA, as follows (23). The result is shown as the MACD line.

$$MACD = 12\, Period\, EMA - 26\, Period\, EMA$$

MACD can gauge whether a security is overbought or oversold, from which traders can obtain signals about a directional move or warning of potential price reversal (23). Traders may buy or sell the security when MACD is below or above its signal line respectively.

The Relative Strength Index (RSI) - another momentum indicator - describes the speed of magnitude of a security's recent price changes. It evaluates overvalued or undervalued securities (24).

$$RSI_{stepone} = 100 - \left[\frac{100}{1 + \frac{Average\, gain}{Average\, loss}} \right]$$

RSI represents a security's strength on days when the price goes up or down related to the strength. When used together with other technical indicators, the RSI can help traders make trading decisions based on a better understanding of the market.

2.2. *Machine learning algorithms*

Decision tree (DT) and random forest (RF) supervised machine learning algorithms were used in this study. In the DT algorithm, a tree-like model is projected hierarchically for making decisions with possible circumstances (25). Then, the best feature is selected to split the data based on the specific criteria at each internal node. This algorithm searches for the parts that provide the most discriminatory data classification power. In contrast, RF is an ensemble learning method to solve regression and classification problems by constructing many small decision trees (25). Multiple decision tree models are provided to make predictions, and individual outputs are combined to enable it to become more accurate with better generalization. The multiple trees are trained on random subsets of the training data, from which individual predictions are made for each tree. The final prediction resulting from the random forest is the overall averaged values predicted by an individual tree. DT is a fast and easy operation approach with large datasets, and it mainly depends on a graphical method to illustrate outcomes with a branching approach. In contrast, RF requires rigorous training, and it consists of multiple sets of DT that take feedbak from the output, providing better accuracy and capability of handling large but previously unmanaged datasets.

A series of the analysis processes applied in this study is illustrated in Figure 3. We used the Google CoLab to write and execute Python code (26). As a first step, the S&P 500 daily stock trend data downloaded from Yahoo Finance (18) was imported into the Google platform. The data was then manipulated to create the trend features (8 indicators) and the corresponding targets (three predictions). Next, the datasets were split into either training or testing samples. Training samples were used for training with machine learning models, while testing samples were exclusively used for evaluating model performances. We used the train_test_split function provided by sklearn (27) to prepare training and testing samples at a 75:25 ratio, which was the default split. After finding the optimal hyper-parameters using 5-fold cross-validation, we trained our machine learning models to optimize model parameters. Testing samples were used to evaluate the model performance using accuracy, precision, recall, and f1-score. We repeated the same procedure for all targets, such as next day, next week, and next month.

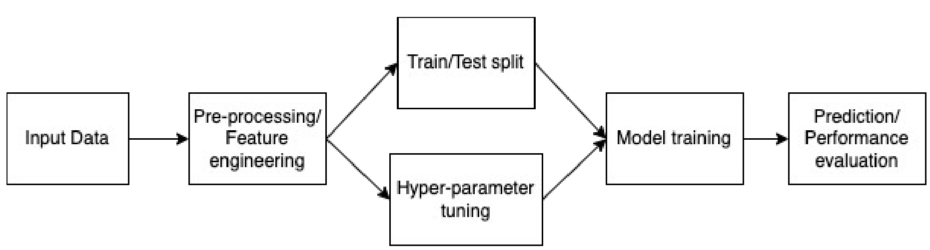

Figure 3. The workflow diagram of the analysis process applied in this study, including pre-data analysis and machine learning algorithm.

For pre-processing, the eight trend features were used as the indicators in the technical analysis, shown in Figures 4 and 6. The full range of 50- and 200-days Simple Moving Average (SMA) adjusted based on the original dataset was compared with the adjusted closing value in Figure 4. It is seen that the 50-days SMA catches up with the adjusted close value, but the 200-days MA shows a smooth transition just behind the peaks and valleys of the 50-days MA. The full range of the calculated RSI (figure 5), MACD, and MACD-signal are also presented with the MACD-signal lagging the MACD. MACD signal was estimated as 9-days EMA of the MACD line. Recent 400-days plots in Figure 6 exhibit detailed trends of the indicators. A MACD below its signal line is usually indicative of a bullish signal, which suggests the price of the assets is likely to experience upward momentum. In contrast, a MACD above its signal line indicates future downward or bearish momentum. RSI usually validates up and down trends, including overbought and oversold securities, providing short-term buy or sell signals.

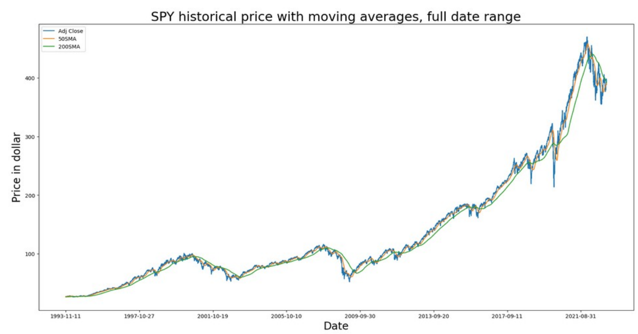

Figure 4. The SPY historical price from 1993-01-29 to 2023-02-17. This plot shows the adjusted close price in dollars and the 50- and 200-day moving average (MA) in dollars from 1993-01-29 to 2023-02-17. The blue, red, and green lines indicate the adjusted close, 50-day MA, and 200-day MA, respectively.

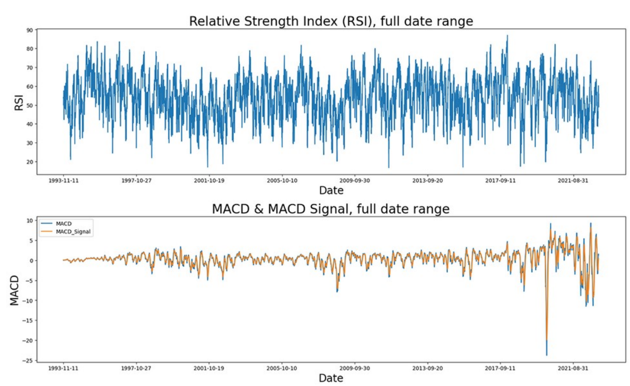

Figure 5. Top) Relative Strength Index (RSI) and bottom) Moving Average Convergence and Divergence (MACD) and MACD-signal calculated for the full range of the input data from 1993-01-29 to 2023-02-17.

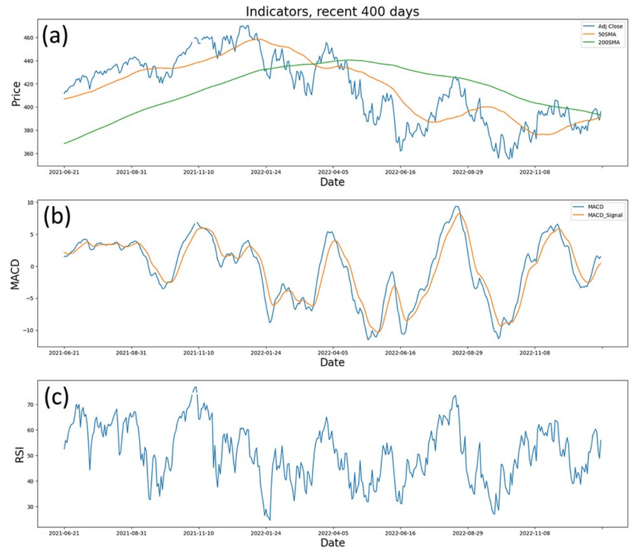

Figure 6. The recent 400 days of the indicators (features) (a) adjusted close volume, 50- and 200-days moving average (MA), (b) Moving Average Convergence and Divergence (MACD) and MACD-signal, and (c) Relative Strength Index (RSI) calculated for the SPY.

3. Results

Accuracy is a metric that evaluates the prediction performance of machine learning algorithms applied (13). The F1 score is a harmonic way of utilizing precision and recall, meaning that both precision and recall contribute equally to the F1 score. The F1 score indicates its best value as 1 for 100% and the worst as 0 for 0%. The F1 score is the harmonic mean of precision and recall. A high F1 score indicates a balanced dataset.The accuracy and F1 score is calculated as follows,

$$Accuracy = Number\ of\ correct\ predictions / Total\ number\ of\ predictions$$

$$F1 = 2 * (precision * recall) / (precision + recall)$$

As discussed above, conventional supervised machine learning algorithms such as decision tree (DT) and random forest (RF) were applied in this study. We made target predictions for the next day, next week, and next month with the accuracy and F1 score, as presented in Table 1. The results show that, overall accuracy and F1 score increased as the target period increased. The F1 score for the next month was ~ 80%, while the accuracy for the next day was ~ 55%. To evaluate the extent of the trending indicators' (features) contribution, we prioritized their contribution in the two applied models, as shown in Figure 7. The y-axis depicts the indicator (features), and the x-axis indicates the contribution percentages of the indicators as a ratio. Thus, if the X value increases, its contribution toward accurate total prediction increases.

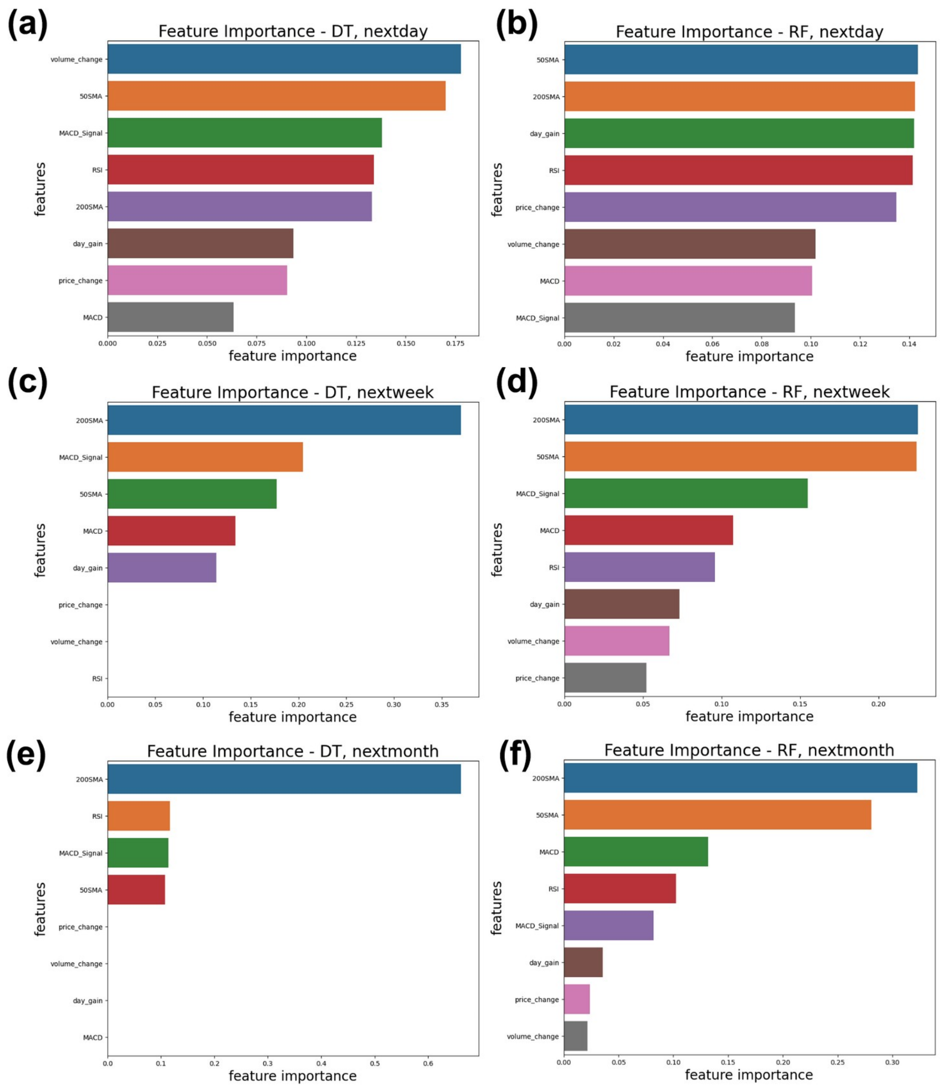

Figure 7. The importance of the features (trend indicators) in the decision tree (a,c,d) and random forest (b, d, f) algorithms for the target projections: next day, next week, and next month, from top to bottom. Each bar in the plot indicates the contribution of features.

It is evident that the volume-change and 50 SMA contribute significantly when using the DT algoritm. For the next week, we were not able to obtain contribution levels of price_change, day_gain, and RSI. For the next month prediction, we could not obtain the level of price_change, day_gain, and MACD. This is because these features contributions were showed no trend for the week and month target extension. However, the 200 SMA demonstrated a significant effect on the next month prediction using the DT algorithm. When using the RF algorithm, the 50 & 200 SMA, day_gain, and RSI demonstrated a similar contribution performance for the next day prediction. All the predictions indicated a preponderant significant domination of the 50 or 200 days SMA. Manipulating the SMA within a specific period was likely most critical to obtaining high predictive accuracy.

The Accuracy and F1 score of the targets: next day, next week, and next month, acquired from decision tree (DT) and random forest (RF) algorithms are presented in Table 1 below.

Table 1. Accuracy and F1 score of the targets: next day, next week, and next month, acquired from decision tree (DT) and random forest (RF) algorithms. Units are in Percent.

	Next Day		Next Week		Next Month	
	Accuracy	F1 score	Accuracy	F1 score	Accuracy	F1 score
Decision Tree	55.7	68.2	59.1	73.4	66.6	79.4
Random Forest	54.4	70.0	58.9	73.9	66.3	79.3

Going beyond the technical approach to predicting the S&P 500 index, we investigated individual stock market prices in several commercial sectors in order to evaluate their performances based only on the previous trends. We investigated our model's predictability for Waste Management, Inc. (WM), Microsoft Corporation (MSFT), JPM Morgan Chase & Co (JSM), and Exxon Mobil Corp (XOM). WM and MSFT are included in the RE100 and are associated with environmental sustainability. WM is a waste management, comprehensive waste, and environmental service company in North America, founded in 1968 (15). MSFT is a multinational software IT company, XOM is an oil and gas energy company, and JPM is a finance and the largest bank by market capitalization in US. In this analysis, we applied the stock trend from 2011-01-03 to 2024-03-14. The market prices of individual stocks are expected to intrinsically differ from the market performance index (S&P 500). However, we hypothesized that our approach would also be able to predict the future stock trend of individual companies in different sectors. RE100 is a global initiative organized by the Climate Group and CDP (Disclosure Insight Action) that comprises more than 400 influential corporations committed to 100% renewable electricity (14). The obtained accuracies and F1 scores for the target projections on the four commercial sectors using the DT and RF algorithms are presented in Table 2. WM is notable in achieving the greatest predicted accuracy. Its stock price predictions for the next month are significantly more accurate when compared to the other individual stocks. Nevertheless, both predicted values for the four cases show a similar trend as that of the S&P 500 analysis. The accuracy and F1 values increased as the target period increased. The overall values were lower than the case of the S&P 500 study. As an example, the contribution levels of the features on the stock price prediction of WM are shown in Figure 8. It is evident that the 50 days SMA is a predominant predictive feature, similar to that seen in Figure 4. This is followed closely by the 200 SMA. The other features show similar contribution levels for the next-day prediction target, but their levels are unequal for the next month prediction target.

Table 2. Accuracy and F1 score (%) of next day, next week, and next month for WM, JPM, MSFT and XOM, using Decision Tree and Random Forest ML processes.

		Next Day		Next Week		Next Month	
		Accuracy	F1 score	Accuracy	F1 score	Accuracy	F1 score
WM	Decision Tree	53.9	67.5	60.0	73.1	68.6	80.0
	Random Forest	56.1	71.4	60.4	74.5	70.4	80.8
JPM	Decision Tree	52.4	66.4	58.9	72.9	64.4	77.8
	Random Forest	52.2	66.0	57.2	72.7	64.7	77.9
MSFT	Decision Tree	53.6	68.8	60.6	74.7	70.6	80.1
	Random Forest	54.8	68.9	60.8	74.8	67.8	79.9
XOM	Decision Tree	50.5	51.0	54.1	69.0	70.5	73.5
	Random Forest	50.5	59.1	58.6	71.3	58.2	71.6

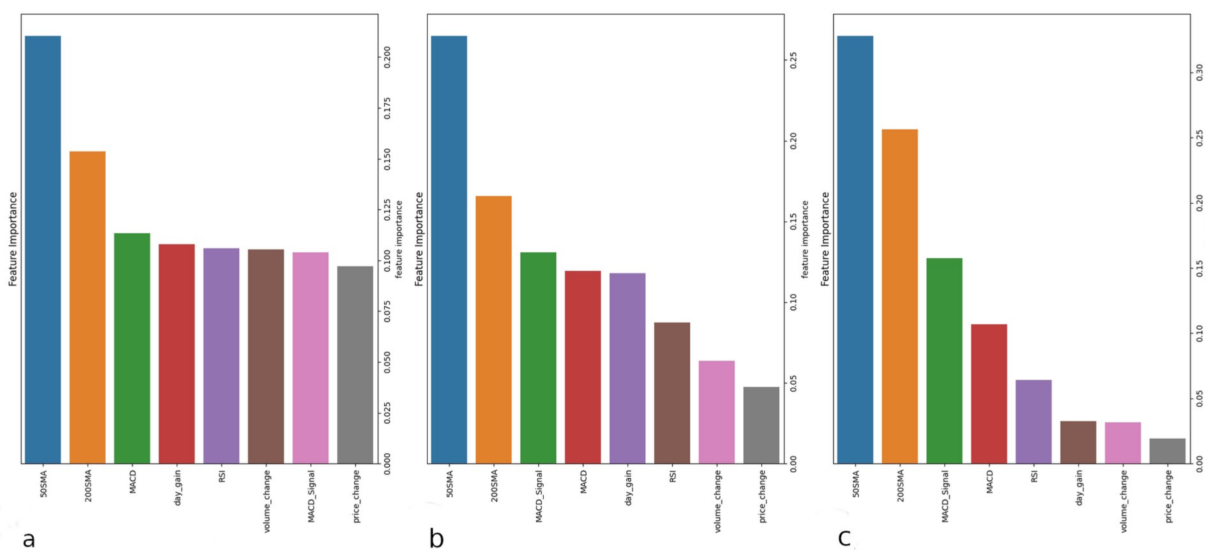

Figure 8. The importance of the features (trend indicators) in random forest (RF) algorithms for the target projections (a) next day, (b) next week, and (c) next month on the stock market (NASDAQ) price of Waste Management, Inc (WM). Each bar in the plot indicates the contribution of features.

4. Discussion

We used technical analysis and machine learning algorithms to predict one-day, one-week and one-month trends for the S&P 500 stock market index. The results showed that increasing the prediction horizon for data training and learning led to greater prediction accuracy. The F1 score for next month target prediction was ~ 80% of the prediction accuracy, while the accuracy for the next day was ~ 55%. A simple moving average (SMA) contributed the most in terms of precise predictions rather than momentum indicators such as MACD, Price, or volume change. Unlike environmental analysis, which is determined by economic circumstances, technical analysis makes direct predictions on future trends using only the evaluation of the previous trends. This type of analysis relies on the trend indicators called features characterized by mathematical or statistical measures. The SMA is obtained by smoothing out stock lines such that it is robust to the short-term and temporary whipsawing up and down movement of the stock price (16). In other words, SMA is time insensitive. This characteristic likely makes the SMA indicator more effective for long-term prediction. In general, the fundamental analysis is considered more suitable for long-term investment, however our study suggests that the predictive capability of this simple indicator or feature using purely technical analysis is comparable to long-term prediction using fundamental analysis.

The Dow Jones/NASDAQ and S&P 500 indexes are among the most widely followed stock market indexes in the US. Stock market. Dow Jones/NASDAQ (National Association of Securities Dealers Automated Quotations) provides automated information about stock prices of companies comprising the information/technology sectors. Different strategic approaches with clear investment goals should be followed to appraise risk tolerance. This is to maximize return with risk minimization. In the same fashion, identical to the S&P 500 study, we predicted the future stock price trends of the four companies belonging to different industrial categories, as described in the previous section. Interestingly, for these individual companies, although overall accuracy values vary when compared with the results from the S&P 500 analysis, the predictive accuracy of simple arithmetic indicators such as the SMA is significantly better when compared with the trend-following momentum and volumes indicators. Likewise, this result could be attributed to the previously mentioned advantage of the SMA.

The Python code used in this study is listed in reference (17). Our study relied on conventional supervised machine learning algorithms such as DT and RF. These algorithms are based on classifying and regressing the original dataset with true or false determination after dividing them into training and learning groups. However, both these algorithms are prone to overfitting. Thus, to minimize this limitation, other supervised algorithms such as Support Vector Machines (SVM) and Naive Bayes (NB) need to be applied to explore hidden structures (if any) of the data groups. In particular, SVM is more powerful for its applicability to high-dimension (or multi-dimension) data. NB is a simple and fast-implemented algorithm. In the future, we plan to further develop our technical approach with these two algorithms for better performance. In addition, we will also add the unsupervised machine learning algorithms to analyze patterns and relationships of the unlabeled datasets without predefining the output.

5. Conclusion

We predicted the future trend of the S&P 500 (Standard and Poor's 500) stock market index, for the next day, next week and next month using machine learning algorithms. We chose price, volume change, day gain, 50- and 200-day Simple Moving Average (SMA), Relative Strength Index (RSI), and Moving Average Convergence Divergence (MACD), as the features. We found that prediction accuracy increased as the defined prediction period increased. Prioritizing the indicators' importance suggested that a Simple Moving Average (SMA) contributed the most to prediction accuracy. Hence, using only technical analysis (and leaving out fundamental analysis), we could effectively predict future index movement based on historical trends using the SMA if the targets were well defined within a period of the movement.

6. Acknowledgement

We would like to acknowledge Dr. Sung for advising and reviewing our technical analysis.

7. References

1. Barro, J R. "The Stock Market and Investment." The Review of Financial Studies, 3(1), 115-131, 1990, https://doi.org/10.1093/rfs/3.1.115

2. "Stock Market", https://en.wikipedia.org/wiki/Stock_market

3. Prasad A, Seetharaman A. "Importance of Machine Learning in Making Investment Decisions in Stock Market." Vikalpa, 46(4), 209-222, 2021, https://doi.org/10.1177/02560909211059992

4. Lee T K, Cho J H, Kwon D S, Sohn S O. "Global stock market investment strategies based on financial network indicators using machine learning techniques." Expert Systems with Applications, 117, 228-242, 2019, https://doi.org/10.1016/j.eswa.2018.09.005

5. S&P500, S&P Dow Jones Indices, https://en.wikipedia.org/wiki/S%26P_500 , https://www.spglobal.com/spdji/en/indices/equity/sp-500/#overview

6. Frino A, Gallagher D. "Tracking S&P 500 Index Funds." Journal of Portfolio Management, 28(1), 44-55, 2001, http://dx.doi.org/10.3905/jpm.2001.319822

7. Prondzinski D, Mitchell M. "Active Versus Passive Investing: Evidence from the 2009-2017 Market." Journal of Accounting and Finance, 18(8), 2018, https://doi.org/10.33423/jaf.v18i8.114

8. Nerdwallet, Active vs Passive Investing: Which is better? https://www.nerdwallet.com/article/investing/active-vs-passive-investing#:~:text=The%20biggest%20difference%20between%20active,of%20investments%20called%20an%20index.

9. Bonga W G. "The Need for Efficient Investment: Fundamental Analysis and Technical Analysis." Social Science Research Network (SSRN), April 11, 2015, http://dx.doi.org/10.2139/ssrn.2593315

10. Fundamental vs Technical Analysis: What is the difference, Indeed.com https://www.indeed.com/career-advice/career-development/fundamental-vs-technical-analysis#:~:text=Fundamental%20analysis%20uses%20economic%20and,future%20price%20of%20an%20investment.

11. Bacher H W, Spindler M. "Machine learning for financial forecasting planning and analysis: recent developments and pitfalls." Digital Finance, 4, 63-88, 2022, https://doi.org/10.1007/s42521-021-00046-2

12. State Street Global Advisors SPDR, SPDR S&P500 ETF Trust, https://www.ssga.com/us/en/intermediary/etfs/funds/spdr-sp-500-etf-trust-spy

13. Dalianis H. "Evaluation Metrics and Evaluation." Clinical Text Mining, Springer, 2018, pp 45-53.

14. Climate Group RE100, https://www.there100.org/

15. Waste Management, https://wm.com

16. Jasemi M, Kimiagari A M. "An investigation of model selection criteria for technical analysis of moving average." Journal of Industrial Engineering International, 8, Article number 5, 2012, https://doi.org/10.1186/2251-712X-8-5

17. https://colab.research.google.com/drive/1ZxJR_VOkmoq_b-ViZFvURmrOMcjlQEDs?authuser=1

18. Yahoo finance, SPDR S&P500 ETF Trust (SPY) https://finance.yahoo.com/quote/SPY?p=SPY

19. Technical Indicator: Definition, Analyst Uses, Types and Examples, Investopedia, https://www.investopedia.com/terms/t/technicalindicator.asp

20. Ozturk M, Toroslu I H, Fidon G. “Heuristic based trading system on Forex data using technical indicator rules.” Applied Soft Computing, 43, 170-186, 2016, https://doi.org/10.1016/j.asoc.2016.01.048.

21. Leigh W, Frohlich C J, Hornik S, Purvis R L. "Trading With a Stock Chart Heuristic." IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 38(1), 93-104, 2008, https://doi.org/10.1109/TSMCA.2007.909508

22. Wikipedia, https://en.wikipedia.org/wiki/Moving-average_model

23. Wikipedia, https://en.wikipedia.org/wiki/MACD

24. Wikipedia, https://en.wikipedia.org/wiki/Relative_strength_index

25. Jehad A, Khan R, Ahmad N. “Random Forest and Decision Tree.” IJCSI International Journal of Computer Science Issue, 9(3), 272-278, 2012, https://www.uetpeshawar.edu.pk/TRP-G/Dr.Nasir-Ahmad-TRP/Journals/2012/Random%20Forests%20and%20Decision%20Trees.pdf

26. Welcome to Colaboratory, Google Colab, https://colab.research.google.com/

27. scikit-learn, Machine Learning in Python, https://scikit-learn.org/stable/