SMU Data Science Review

Volume 7 | Number 1 Article 6

Profiting from Dow Jones Industrial Index and Hang Seng Index using moving average and MACD optimization model

Anthony Yeung
Southern Methodist University, ayeung@smu.edu

Joe Wailun Chung
Southern Methodist University, joechung@smu.edu

Nibhrat Lohia
Southern Methodist University, nlohia@smu.edu

Onyeka Emmanuel
Southern Methodist University, oemmanuel.datascience@gmail.com

Profiting from Dow Jones Industrial Index and Hang Seng Index using moving average and MACD optimization model

Anthony Yeung[1], Emmanuel Onyeka, Joe Wailun Chung, Nibhrat Lohia
[1] Master of Science in Data Science, Southern Methodist University, Dallas, TX 75275 USA
Ayeung@smu.edu, oemmanuel@smu.edu, joechung@smu.edu

Abstract. Before the internet, high-speed laptop computers, and big data became accessible and popular, academia on stock market trading concentrated on Efficient Market Hypothesis (EMH). EMH hinges on the idea that the market is efficient and there is no extra return that could be generated. With the dynamic development of the internet, big-data and computing technology, many researchers started to pay attention to Technical Analysis and its usage. Numerous academic papers claimed that technical analysis can enhance returns by using various technical tools. This paper explores in-depth the simulation model of Moving Average and Moving Average Convergence/Divergence (MACD) to come up with optimized parameters that will allow traders to profit from trading Dow Jones Industrial Index and Hang Seng Index.

1 Introduction

Investing in stocks is a great way to build wealth. A lot of individuals have successfully become millionaires and even stock market whales buy making smart investments. Such life changing returns takes time, decades even, but the benefits of stock trading make it worth the wait. One such benefit is that anyone can get into stock trading. One does not even need a lot of money to start trading stocks. A trader can start small, investing what little they have for the time being, and incrementally increasing their investments as their financial abilities grow. Another benefit of stock trading is that it can be a source of passive income. Though the effects of this income may not be felt, as gains are re-invested to produce even higher gains, the money is there and ready when needed. Given these great benefits of stock trading, there are some challenges that cannot be overlooked.

One such challenge is navigating the market. The American stock market is huge and can be confusing to newcomers. Some seasoned stock traders even find themselves needing professional advice from time to time. Even with the vast wealth of books, articles, videos, seminars, and experience that are available, traversing the financial market while avoiding its land mines and pitfalls is very challenging. Not many people have the patience, time, and willingness to thrive in stock trading. A lot of times one might make a trade simply on a hunch as if stock trading is the digital version of playing the lottery or a virtual casino where luck is paramount. The slightly wiser ones admit they do not know what they are doing and so they base their trades on the results they see after a quick google search of the top 10 stocks to invest in today. If financially capable, some may even invest in hiring someone or an investment firm to handle their portfolio. If only it were that easy. Just pay a firm to invest on one's behalf and reap the sweet rewards of their labor.

This approach is likely only useful if the trader has a high amount of capital to begin with. The trader would have to make investments in the thousands to experience returns that are high enough, and early enough for it to matter. Of course, there is short-term trading, but this is riskier, and requires more knowledge of how the market works. Since an investment firm does the research, there are fees that must be paid to them which reduces overall gains. Overall, this is not a bad strategy, and it is highly recommended. However, one can learn the trade, cut out the middleman, and maximize returns.

Learning to read trade signals may look like rocket science, but it is not difficult at all. Research in this area has been ongoing, with many researchers and investors having made great contributions to successfully minimizing loss while increasing profits. Besides the ongoing research, there are also better, more reliable tools out there that help analyze the market. People are building trading bots that analyze the market and push trade signals to the buyer. Some of these bots even go as far as to do the trading based on set parameters by the trader.

The great thing about such bots is that they are easily accessible and do not cost an arm and a leg to get. Another great thing is that they are quicker than humans at initiating a trade. This is especially useful if the market is flooded. Often a trader can enter an order for a stock, but by the time it goes through the stock has risen. This is especially noticeable when day-trading. Also, they are likely to catch quick changes in market price trend and make the necessary trades.

Like with humans however, trading bots are not infallible. Technical or software failures can occur, and they can make bad trading decisions. The bot user may not know the algorithm or strategy being used by the bot. Trading bot services offer default strategies the users can choose from, but these strategies may not be good. There is the possibility of creating custom strategies for the trading bot. However, unless one knows how to program, a programmer will likely need to be hired to do this. Trading bots require maintenance and they do not completely absorb the user of looking at the market from time to time. Also, though using trading bots is completely legal, some trading platforms do not allow their use. This could have a limiting effect for the trader as not all trading platforms are the same. They may differ in their rules and policies, minimum investment requirements, and requirements to be a day trader, and even stocks that are tradable on the platform.

Regardless of if the trade is being made by a trading bot or by a human, trade decisions must be backed up by data. Some people have employed the use of technical analysis to help understand market trends. Such analysis helps detect patterns in the stock price over time and correlate those patterns with other patterns to help predict upcoming trend changes. Such people use advanced machine learning methods to help navigate the stock market. Methods such as Neural Networks have been used to optimize stock trading models and achieve more accurate trading signals and less volatile trends. Despite these advancements in the field, there are still several issues that make it challenging to engage and beat the financial market.

The first issue is volatility. The stock market is extremely volatile and unpredictable by normal means. The trend in the price of stocks is determined by how many investors are selling or buying during a period. When investors sell, the stock price goes on a downward trend, and when they buy, it goes on an upward trend. This behavior makes the market difficult to predict, and results in a wide spectrum of fluctuations. To get ahead of these trends, one would have to be familiar with factors that tend to affect investors to buy or sell and be able to monitor such factors.

Another issue is that of monitoring the market without missing details. Unless one is a full-time stock market trader or is working with a brokerage firm, there is only so much time that can be given to monitoring the stock market. Having a system that does this and automatically makes trade suggestions would be a more effective approach.

Ultimately, the goal for anyone attempting to engage the stock market is to establish a pattern of consistent gains. Though every investor generally has this goal, their approach towards achieving this goal may vary. These different approaches vary in ease of use, performance and knowing which approach to take is another challenge that investors must face.

The aim of this research is to address these issues and challenges by working out a profitable, simple to understand simulation model for the Dow Jones Industrial Index and Hang Seng Index. The goal is to find out the best parameters to use in Moving Average and Moving Averages Convergence/Divergence (MACD) models. Simulation methodology will be done using Python programming language. Moving Average and Moving Averages Convergence/Divergence (MACD) are the focus because they are simple and easy to understand as there is no complex math involved. Time Series Momentum (TSMOM) will be employed using moving average trading rules to improve the MACD model.

2 Literature Review

The literature review focuses on four main areas or sections. The first section will cover some key concepts and factors that are used to interpret the stock market. The second section will be a direct comparison between fundamental analysis and technical analysis, showing some of their pros and cons. The third section will focus on what the Efficient Market Hypothesis is, and how effective fundamental and technical analysis are according to this theory. The fourth section will compare the buy-and-hold strategy with that of evaluating moving averages and evaluate which performs better.

2.1 Divergence vs. Convergence

Interpreting visual representations of any stock market price is about recognizing patterns and observing correlations in the stock price trend to those patterns. One can explain what patterns to look for, but ultimately there are foundational things that must be understood to understand these patterns. Pattern recognition is a common practice in stock market trading, and it is not going away any time soon. Usually, patterns in the stock price are looked for, but it does not end there. Patterns in the technical analysis indicator, such as MACD, are observed along with stock price patterns to detect any new meaning in the price. This also helps with increasing the predictability of trends and the accuracy of the timing of trend shifts. There are two typical patterns that are watched for. The first is divergence, and the second is convergence.

Divergence is when the stock price has a different trend from the indicator (Adrian et al., 2011). There are two types, classic or regular, and hidden (Martins et al., 2017). Regular divergence is when there is a significant upward or downward trend in the stock price, but the same trend is not observed in the indicator (Hung et al., 2016). For example, if the stock price is on a downward trend, and the MACD line is flat or on an upward trend this could be an indication that the stock price is about to reverse its trend.

This is an example of classic divergence. In classic divergence the trend of the indicator, otherwise known as momentum, should agree with the trend of the stock price (Martins et al., 2017). If classic divergence is being observed, it is an indication to get ready to buy or sell as the market will soon be reversed.

Other factors however may come into consideration. One such factor is whether the trend is bearish or bullish. Bearish trends will have less of a slope than bullish trends. Another factor is how quickly the market reverses after a divergence is observed. Sometimes after a divergence, the market reverses, but does not get back to a level position. A level position is a period during the divergence where the stock price was neither on an upward or downward trend. Because of these two factors, regular divergence buy itself should not be used as an indicator for predicting if one should buy or sell.

To complement regular divergence, hidden divergence should be used. Hidden divergence is when there is a significant upward or downward trend in the indicator, but the same trend is not shown in the stock price (Martins et al., 2017). Whenever hidden divergence is observed, this is an indication that the trend seen in stock prices will likely continue as is. Unlike regular divergence which is usually observed on the tail end of a momentum, hidden divergence is usually observed before the tail end.

Convergence is more straightforward than divergence. It is when the stock price and the technical indicators have the same trend or direction (Martins et al., 2017). Just like divergence, it can be bearish or bullish. It can also be a negative or positive convergence. In a negative convergence, the trend of both the price and technical indicators are going downward. A positive convergence shows both trends going in an upward direction. Convergence also differs from divergence in what it is used for.

Convergence is used as a strong indicator that the observed trend is not a fluke. Unlike divergence, convergence is typically not used to gauge if a trader should buy or sell (Martins et al., 2017). It does not predict an upcoming change in the observed trend. Rather, it merely serves as a reassurance that the trend will hold. So, for example, if a downward convergence is observed, the trader knows to wait as prices will likely keep going lower. Similarly, if it was an upward convergence, this would be reassurance for the trader to hold their positions and not sell.

The goal of looking for patterns of divergence and convergence is to determine if a stock price has been overbought or oversold. When a stock price has been overbought, it will be on an upward trend and will soon switch to a downward trend (Adrian et al., 2011). When it is oversold, the trend is usually downward, but will soon switch to an upward trend (Adrian et al., 2011). Divergence can happen quickly so it must be monitored very frequently (Adrian et al., 2011). This sounds simple, but there are some issues that may arise.

One such issue is that divergence and convergence do not tell how far the trend will go. For example, observing a regular divergence pattern may indicate that a switch in trend will soon occur, however it does not indicate how long the trend will continue after the switch. A lot of times what happens is that one will buy or sell because this pattern was observed, and the trend switches for a short period, could even be an hour, then reverses back and falls even further. Even if this was not the case, there is no guarantee that when the trend is switched, it will rise past its previous support point, the point where it was last level or flat. It is important to note that ultimately, the goal is to make the best educated guess, and hope for the best.

2.2 Fundamental Analysis vs. Technical Analysis

Finding which company's shares to buy can often be as challenging as trying to find a needle in a haystack. Given the high number of companies that have flooded the financial market, it is imperative that one takes time to evaluate which ones are worth paying attention to. There are two major forms of analysis investors use to analyze the financial market (Suresh et al., 2013), and thankfully, one of them, fundamental analysis, attempts to address this very issue.

Fundamental analysis evaluates underlying factors that affect or could potentially affect the economy, financial market, and company (Suresh et al., 2013). In other words, this type of analysis is used to measure the quality or intrinsic value of a stock. It hinges on the assumption that the performance of the securities that the company is representing is a direct or indirect reflection of said company (Suresh et al., 2013).

The good thing about fundamental analysis is that it uses a bird's eye view. As such, one can make broad generalizations of a company by evaluating things such as the economy in general, the industry, and even the political state nationally or internationally (Suresh et al., 2013). Nothing is off the table with this approach. A con of this approach is that a lot of the results and data are not objective and thus are open to potentially faulty interpretation (Suresh et al., 2013).

Despite this, there are several things that fundamental analysis struggles to determine for an investor. First is when to invest, and next is when to sell one's shares. In the financial world, timing matters as prices usually fluctuate between an upward or downward trend. Unless one intends to invest for the long haul, selling on a downward trend is crucial to make profit. As such an analysis of the price trends will help address this issue. Thankfully, the second of the two major forms of financial analysis can be used to address this issue.

Technical analysis is the second of the main schools of thought for financial market analysis. Unlike fundamental analysis, technical analysis focuses on price movement of a security and attempts to predict future price movements using that information (Suresh et al., 2013). In other words, this type of analysis primarily looks at historical prices to identify optimum trading opportunities rather than evaluate the quality of the stock (Suresh et al., 2013). It focuses on providing a clear picture of the financial patterns of a company.

What's great about technical analysis is that it appears to take less time and effort than fundamental analysis (Abarbanell et al., 1997). This is because price data is usually more readily available and analyzing this data can easily be done using computer programs. Another great thing is that the information evaluated using technical analysis is more objective and less open to interpretation, as opposed to fundamental analysis (Abarbanell et al., 1997). This, however, means that the scope of this analysis is much narrower. Due to the specific nature of technical analysis, there is often the risk of including too many indicators or signals (Abarbanell et al., 1997). Also, using this approach by itself does not consider mitigating factors that could be great predictors of share price.

Upon understanding both fundamental analysis and technical analysis, it seems that each one is complimentary to the other. It makes sense that to get a wholesome picture of the financial market, one needs to not only evaluate the trends of its companies, but also consider the underlying factors that are affecting those companies. Despite all this, there is the question of how reliable these two forms of analysis really are. One must wonder if they are useful for long-term investments, or just for short-term gains. Are the benefits of fundamental analysis and technical analysis significant enough to be considered proven methods for establishing consistent gains in the financial market? To answer this question, there is one theory that must be looked at, the Efficient Market Hypothesis.

2.3 Efficient Market Hypothesis (EMH)

One commonly held theory for investing in the stock market is the Efficient Market Hypothesis. This theory states that the share price of a stock already factors in all publicly available information (Michael et al., 1989). In other words, there is no information online or otherwise that one can find to help them beat the market (Degutis et al., 2014). This theory implies that the only way to beat the market is to find information that is not publicly available as that information would not already be factored into the share price. This is not to say that one cannot occasionally beat the market, but that one cannot consistently beat the market, and that in the long run, the randomness of market share prices makes it impossible to be predicted (Degutis et al., 2014).

EMH categorizes the efficiency of the financial market into three forms, weak, semi-strong, and strong. These three forms primarily differ in the information that the share price reflects (Michael et al., 1989). The weak form of EMH suggests that the share price reflects all previous or historical prices as securities (Michael et al., 1989). Since all past prices are already factored in, technical analysis, which primarily evaluates historical prices, is rendered useless (Degutis et al., 2014). Fundamental analysis can however provide short-term returns as this kind of analysis evaluates various economic and financial factors that affect a business (Michael et al., 1989). Long-term returns should not be expected as there is still a high level of randomness in this form (Degutis et al., 2014).

In its semi-strong form, the share price reflects all past prices, and publicly available information, but it does not contain inside or private information (Michael et al., 1989). In this form, both technical and fundamental analysis are useless as all new public information is quickly priced into securities. Unlike the weak form, in this form, one can hope to beat the market using inside information. However, one problem with this is that inside information that can help one beat the market is usually known only to those within the organization selling the shares (Fishe et al., 2004). Even if one were fortunate to work within the organization, using such information or leaking it to third parties is illegal (Fishe et al., 2004). Also, it has been shown that inside trading has a negative impact on market liquidity (Fishe et al., 2004).

In the strong form of market efficiency, all past prices, publicly available information, and inside information is already reflected in the share price (Michael et al., 1989). In this form, the investor stands no chance of beating the market, even with inside information (Michael et al., 1989). In this form, as well as the previous two forms, one can still get abnormal returns as there are anomalies or outliers that will happen, but the operative word here is consistency. Establishing a pattern of consistent returns is impossible with this form (Degutis et al., 2014).

In summary, according to EMH, the two main schools of thought for financial analysis, technical and fundamental analysis, are mostly useless and inefficient for establishing consistent returns. Also, ethical factors such as this can dissuade potential investors from engaging in the financial market using inside information. It is therefore important to ask three questions. The first question is if EMH is correct. The second question is if there are other strategies that have been proven to yield consistent returns. Finally, there is the question of if machine learning techniques can help yield consistent returns.

To answer the first question, though there are critics of EMH, the theory is still true today (Malkiel et al., 2005). For the most part, investors in both U.S. and foreign markets, do not outperform their index benchmarks, and market prices appear to reflect all available information (Malkiel et al., 2005). Before giving up on the goal of beating the financial market consistently, one must look to see if there are any strategies that have been proven to establish consistent gains.

2.4 Buy-and-Hold vs. Moving Average Convergence/Divergence (MACD)

Two other strategies that have stood the test of time are the buy and hold strategy, and the Moving Averages Convergence/Divergence (MACD) approach. The buy and hold strategy, as the name suggests, is one in which an investor buys shares and does not sell them for a long period of time (Yen et al., 2010). With this approach, one ignores market price fluctuations, thus making this a more passive approach (Yen et al., 2010). The danger of this approach is that one may be forced to sell at a less than optimal time (Yen et al., 2010).

MACD is as much an indicator as it is a strategy (Yazdi et al., 2013), and as the name suggests, it works using the average over multiple periods of time (Yazdi et al., 2013). For example, one can have the moving average for a 20-day period, and then for a 15-day period, then subtract the latter from the former to get the moving average plot. This is usually used in tandem with an exponential moving average which tends to cut between the wave like plot of moving average previously calculated (Yazdi et al., 2013).

MACD is useful for identifying exact turning points in the financial market (Yazdi et al., 2013). In other words, it is used to see when the market will shift from upward trend to downward trend and vice versa. Whenever the share price or stock goes below the moving average, this indicates that the price will continue to go downward, and when it goes up and crosses the moving average, this indicates that the price will continue to rise. As such, unlike with buy-and-hold strategy, MACD requires an active eye on the market and is used for more short-term trading (Sobreiro et al., 2016).

Though buy-and-hold strategy historically has performed satisfactorily, it has been shown to be outperformed by various other strategies (Yen et al., 2010). For example, pitting the test for superior predictive ability (SPA) against the typical buy-and-hold strategy showed that nine out of ten times, market efficiency remained consistent and outperformed the buy-and-hold strategy (Yen et al., 2010). In another study, moving average strategies were employed on a portfolio composed of stocks with a high probability of informed trading, and the returns were significantly higher than when buy-and-hold strategies were employed on the same portfolio (Hung et al., 2021). Also, A recent study examined if variable and fixed-length moving averages (VMA and FMA) could outperform the typical buy-and-hold approach in the Pakistan stock market. The results suggested that VMA and FMA strategies are generally more profitable than buy-and-hold, and that they are consistently profitable so traders in the Pakistan stock market can use these strategies reliably (Khand et al., 2020).

Unlike the buy-and-hold approach, the MACD approach has been shown to be high performing, and often outperforming its counterparts. This study compares the MACD approach to the buy-and-hold approach while using The Genetic Algorithms strategy and the technical analysis strategy to determine which yields higher profit margins. The Genetic Algorithms outperformed the technical analysis strategy by around 4 % (Aguirre et al., 2020). Also, the MACD approach was shown to outperform the buy-and-hold approach (Aguirre et al., 2020). Another study which looked at how profitable the index trading strategies based on dual moving average crossover (DMAC) rules were in the Russian stock market showed that DMAC strategies typically outperform buy-and-hold strategies, though not by a significant amount (Luukka et al., 2016). In another study, the Moving Average Convergence Divergence (MACD) and Relative Strength Index (RSI) were compared, using six decades of data, to see which approach is more profitable. MACD significantly outperformed RSI using the rules of profitability (Chong et al., 2008).

Despite the superiority of MACD, there are some other factors that one needs to be aware of that affect its performance. One such factor is geo location. Technical analysis shows that relative to the buy-and-hold strategy, the performance of MACD varies from country to country (Sobreiro et al., 2015). Some countries experience better performance from the moving average approach while some don't (Sobreiro et al., 2015). This result is further supported by another study on the performance of the Simple Moving Average, Exponential Moving Average, MACD, and Triple Screen, which showed that these techniques yield returns that are higher than the investment value but have low predictability power in the Brazilian market (Costa et al., 2015).

Another such factor is trading costs. These could include transaction costs, commissions, brokage charges, taxes, etc. Though this is not as significant as geo location, it is worth mentioning that this must be taking into consideration when gaging the performance of a trading strategy (Avramov et al., 2020). The gain from an approach must cover these costs otherwise one would incur a loss. In one study, it was shown that long-run moving averages of prices (MAD) are great at predicting future equity returns, and that these returns are not negated by trading costs that tend to be accrued by institutions (Avramov et al., 2020).

Hopefully, in this study, an optimized trading strategy that will result in a win-rate that is 70% or better will be built. It should beat the buy-and-hold strategy, and it should not have a big drawback which negatively impacts the stability of the model. For this the hope is that the strongest technical indicators will be identified and used in tangent with the stock price data.

3 Methods

3.1 Data

This study focuses on the Dow Jones Industrial Average (DJIA) Index and Hang Seng Index (HSI). 15 years of daily data is collected from Yahoo finance on over 500 stocks. The data includes the Date, Time, Open, High, Low, Close, Volume, and Turnover since 2008. This data source is Yahoo Finance. Wikipedia has a list of all the stocks contained on Yahoo Finance. This list will be imported for the back testing.

3.2 Comparing Models

A few models will be used for this research, and though they all have the same goal of maximizing the win rate, and profits or return, the approach used in each of them will vary. Also, all models will be momentum based. The first model will be the MA model which will use Simple Moving Averages (SMA) of the stock prices over a select number of days. The second model will be the Moving Average Convergence Divergence (MACD) model, which will use the Exponential Moving Averages (EMA) of the stock prices. EMA is just like SMA except that the more recent the data is, the more weight it carries in the model. Each model will have its parameters tuned through simulation and iteration of multiple parameters to find which ones will give the best results.

Other features will be factored into the final model. For example, a huge mistake traders make is to trade during periods when the price is on a downward trend. For this reason, the optimized trading model will have a built-in safety feature that will not permit trading if the market price is on a down trend. This will minimize losses and prevent overtrading. There may be exceptions to this, but further testing and analysis will determine what these exceptions will be, if any. Also, trading costs and other associated fees will be factored into the models.

Moving Average (MA)

For the MA model, a crossover between MA_1 and MA_2 will be used for the buy and sell signals. The variable, MA_1, is the price itself or moving average of the price. It is the short line of the trade. The variable, MA_2, is the longer line of trade. For example, if MA_2 is the moving average calculated for a period of 10 days, then it is the simply the mean of the price for the past 10 days. MA_1 will be the mean of the price for any number of days that is less than what is specified for MA_2. For buy signal, MA_1 should cut above MA_2 from below. For sell signal, MA_1 should cut below MA_2 from above. Parameters will be tested to avoid false signaling.

The optimal parameters of MA_1 and MA_2 should achieve maximum profit during the period being examined. To do so, multiple parameters will be simulated to find out the optimal values. Several versions of this model will be created and tested. Though both will be returning the same thing, they will vary in approach. It is expected that of all the models that will be tested, this should be the most straight forward.

The dates where the short MA crosses above the long MA, and the dates where it crosses below the long MA will be calculated. Then it will be translated to buy/sell signals, when MA_1 crosses above MA_2, then the variable is set to the buy signal. If MA_1 crosses below MA_2, then the variable is set to the sell signal. The return is calculated according to the closing price of the buy/sell dates. In addition, the win rate percentage is then calculated.

$$V = price \tag{1}$$

$$MA_t = \frac{V_{t-n+1} + V_{t-n+2} + \cdots + V_t}{n} \tag{2}$$

Moving Average Convergence/Divergence (MACD)

Like the MA model, MACD is a calculation of the moving average, but unlike the MA model, it is calculated by the Exponential Moving Average (EMA). Traditionally, 26-days, 12-days, and 9-days periods are used. Mechanical rules will apply using MACD_1. These rules will not allow trading if MACD_1 is below 0. MACD_1 above 0 indicates an upward trend, while MACD_1 below 0 indicates a downward trend. Besides the use of EMA, this model will function like the MA model.

$$n = days\ period \tag{3}$$

$$EMA_t = V_t \times \left(\frac{2}{1+n}\right) + EMA_{t-1} \times \left(1 - \frac{2}{1+n}\right) \tag{4}$$

$$MACD = short\ period\ EMA - long\ period\ EMA \tag{5}$$

Fig 1. MACD model with fast – 12 and slow – 26 for DJIA between 2017 and 2019

3.3 Metric

Several metrics will be used to measure model performance. The first is profits, which will be calculated by dividing the difference between the selling price and the buying price by the buying price.

$$\Pr o\ fits\ =\ \sum \frac{(Selling\ price - Buying\ price)}{Buying\ price} \tag{6}$$

It is expected that some of the profits may be negative as it is possible to lose money. As such the model will track losses and treat it as importantly as profits. Though maximizing profits, and minimizing loss are two different things, they will both be assigned weights that will be used to determine which to focus more on, and when. This will be a risk versus safety decision where taking risks will serve to maximize profits and playing it safe will minimize loss.

Another metric that will be used to measure model performance is win-rate. Win-rate is calculated by dividing the number of wins by the number of profits. This metric is useful to gauge if the model is optimized more for long-term or short-term trading. It is expected that it is possible to get a perfect win rate. A perfect win-rate will be avoided as it is unrealistic and will likely be an indication of a bad model or an outlier. A win-rate of over 0.7 will be acceptable though the goal is to go as high as possible.

$$Win\ Rate = \frac{number\ of\ wins}{number\ of\ wins + number\ of\ losses} \tag{7}$$

Though profit and win-rate will be the primary metrics used, other things will be factored in. One such thing will be how long the trader is willing to wait to experience these profits. Also, a model that produces a high number of perfect win-rates will be unusable. The amount lost will be another metric that will be looked at.

Though this model will be great for short-term trading, as it is designed to have a maximum holding of 10 days for the stock, it will be adapted for long-term trading. This will be done through manipulation of the RSI, number of days for the MA, and a wider date range. The ideal version of this model will be able to perform well for both short-term and long-term trading. If this is unattainable, it will be tuned to focus on either one.

3.4 Cross Validation

Financial markets are evolving, one set of fixed parameters may not perform well all the time. Cross-validation is used to prevent over-fitting when finding the optimized parameters for the models. 15 years of data are used and split into 5 different sets on a rolling basis.

For each model, the optimal parameters are found using the control set. Then, the parameters are applied to the test set. After 4-fold validation, the model with the highest average return from the test sets is selected.

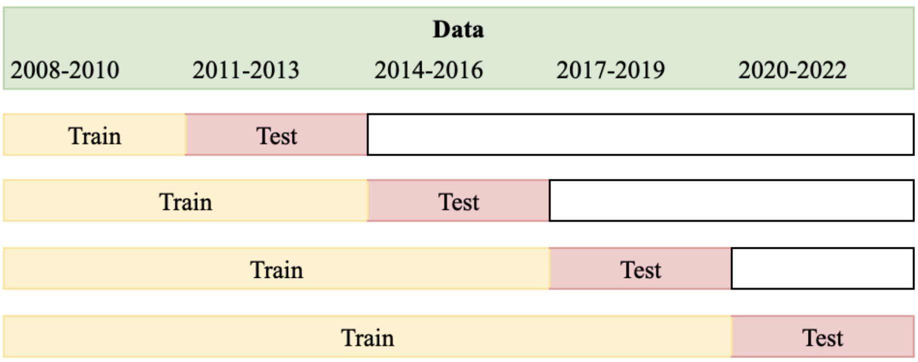

Fig 2. Validation Sets

4 Results

4.1 Results

Since the typical period of days used to calculate the fast-moving average and slow-moving average is 12 and 26 respectively, these values were used as a benchmark against randomly chosen fast- and slow-moving average values. Each run yielded interesting varying results. Also, the models were split into 4 groups which varied in the number of years of data used in the training model. The test groups in each of these 4 groups were fixed to 3 years of data as seen in Table 1. The MACD and MA approaches were applied for each group. Also, the fast- and slow-moving averages for each approach were calculated for a random number of days. These number of days were divided into 2 groups. The first group was limited to 30 days, and the second group was limited to 60 days.

Table 1. Model groups and Test – Train Splits

Group	Approach	MA (#DAYS)	Training Data (years)	Test Data (years)	Test-Train Split (%)
Group A	MACD	Benchmark	3	3	N/A
		Max 30	3	3	50
		Max 60	3	3	50
	MA	Benchmark	3	3	N/A
		Max 30	3	3	50
		Max 60	3	3	50
Group B	MACD	Benchmark	6	3	N/A
		Max 30	6	3	33
		Max 60	6	3	33
	MA	Benchmark	6	3	N/A
		Max 30	6	3	33
		Max 60	6	3	33
Group C	MACD	Benchmark	9	3	N/A
		Max 30	9	3	25
		Max 60	9	3	25
	MA	Benchmark	9	3	N/A
		Max 30	9	3	25
		Max 60	9	3	25
Group D	MACD	Benchmark	12	3	N/A
		Max 30	12	3	20
		Max 60	12	3	20
	MA	Benchmark	12	3	N/A
		Max 30	12	3	20
		Max 60	12	3	20

Looking at Figure 2, there are several interesting things to point out. The average returns seem to be significantly higher with the DJIA Index than with the HSI Index (p-value = 0.01). Though this difference exists, when measuring model performance using win rate percentage it appears not to be significant (p-value = 0.36). This lack of significance can also be seen in the difference between the means. The average win rate percent for models using the DJIA Index is less than three points higher than that of models when using the HSI Index. The same is 14 points higher when measuring with the average return percentage.

Return % DJIA vs HSI

Welch Unpaired T-test		
MEAN	14.5	5.5
SD	12.09	9.53
SEM	2.7	2.13
N	20	20

t score	df	standard error
2.6142	36	3.443

Confidence Int	Confidence %	2 tailed P-val
2.02 - 15.98	95	**0.01**

Win Rate % DJIA vs HSI

Welch Unpaired T-test		
MEAN	55.15	52.65
SD	10.39	6.34
SEM	2.32	1.42
N	20	20

t score	df	standard error
0.9188	31	2.721

Confidence Int	Confidence %	2 tailed P-val
-3.05 - 8.05	95	**0.36**

Fig 3. T-Test of Average Return % & Win Rate % of DJIA vs HSI

This difference in return percentage seems to remain insignificant regardless of if the model is a MACD model or if it is a MA model. The same is true for the difference in win rate percentage. When looking at Figure 3, the difference in return and win rate percentage between models using the DJIA Index and the HSI models is shown to not be significant for MACD models (p-value = 0.08). This also appears to be true for the MA models (p-value = 0.3) as seen in Figure 4.

Return % DJIA vs HSI (MACD Only)

WELCH UNPAIRED T TEST		
MEAN	18.88	5.63
SD	13.08	14.53
SEM	4.62	5.14
N	8	8

t score	df	standard error
1.9172	13	6.911

Confidence Int	Confidence %	2 tailed P-val
-1.68 - 28.18	95	**0.08**

Win Rate % DJIA vs HSI (MACD Only)

WELCH UNPAIRED T TEST		
MEAN	59.13	53
SD	11.34	7.84
SEM	4.01	2.77
N	8	8

t score	df	standard error
1.2564	12	4.875

Confidence Int	Confidence %	2 tailed P-val
-4.50 - 16.75	95	**0.23**

Fig 4. T-Test of Avg Return % & Win Rate % of DJIA vs HSI Indexes for MACD Models

Return % DJIA vs HSI (MA Only)

WELCH UNPAIRED T TEST		
MEAN	7.13	3.25
SD	9.01	4.23
SEM	3.19	1.5
N	8	8

t score	df	standard error
1.1004	9	3.521

Confidence Int	Confidence %	2 tailed P-val
-4.09 - 11.84	95	**0.3**

Win Rate % DJIA vs HSI (MA Only)

WELCH UNPAIRED T TEST		
MEAN	50.25	52.5
SD	8.4	6.07
SEM	2.97	2.15
N	8	8

t score	df	standard error
0.6142	12	3.663

Confidence Int	Confidence %	2 tailed P-val
-10.23 - 5.73	95	**0.55**

Fig 5 T-Test of Avg Return % & Win Rate % of DJIA vs HSI Indexes for MA Models

On average, all the MACD models performed better than the MA models. This outcome was observed regardless of the slow and fast MA values used in the model. Also, this was the case in both the DJIA and HSI Indexes, and it was true regardless of whether the average win rate or return was used to measure model performance. The difference in performance between the MACD and MA models however appears to only be significant when using the DJIA Index (p-value = 0.05) as seen in Figure 5. For models using the HSI Index, the difference in average return percentage between MACD and MA models is not significant (p-value = 0.67).

Return % MACD vs MA (DJIA Only)

WELCH UNPAIRED T TEST		
MEAN	18.88	7.13
SD	13.08	9.01
SEM	4.62	3.19
N	8	8

t score	df	standard error
2.0925	14	5.615

Confidence Int	Confidence %	2 tailed P-val
-0.48 - 23.98	95	**0.05**

Return % MACD vs MA (HSI Only)

WELCH UNPAIRED T TEST		
MEAN	5.63	3.25
SD	14.53	4.23
SEM	5.14	1.5
N	8	8

t score	df	standard error
0.4439	8	5.351

Confidence Int	Confidence %	2 tailed P-val
-9.96 - 14.71	95	**0.67**

Fig 6. T-Test of Avg Return % of MACD vs MA models for DJIA & HSI Indexes

Win Rate % MACD vs MA (DJIA Only)

WELCH UNPAIRED T TEST		
MEAN	59.13	50.25
SD	11.34	8.4
SEM	4.01	2.97
N	8	8

t score	df	standard error
2.0925	14	5.615

Confidence Int	Confidence %	2 tailed P-val
-0.29 - 23.79	95	**0.05**

Win Rate % MACD vs MA (HSI Only)

WELCH UNPAIRED T TEST		
MEAN	53	52.5
SD	7.84	6.07
SEM	2.77	2.15
N	8	8

t score	df	standard error
0.1426	8	3.505

Confidence Int	Confidence %	2 tailed P-val
-7.07 - 8.07	95	**0.89**

Fig 7 T-Test of Win Rate % of MACD vs MA models for DJIA & HSI Indexes

index	fast_slow_range	model	group	fast	slow	win rate %	return	return %
SORTED BY WIN RATE % FOR TEST GROUPS (TOP 5)								
DJIA	max 30	MACD	C	8	18	0.72	35860.82	0.35
DJIA	max 30	MACD	A	13	26	0.66	25812.65	0.25
DJIA	max 60	MACD	C	5	47	0.65	36395.36	0.36
DJIA	benchmark	MACD	A	12	26	0.64	24756.41	0.24
DJIA	benchmark	MACD	D	12	26	0.64	19825.14	0.19
SORTED BY RETURN % FOR TEST GROUPS (TOP 5)								
DJIA	max 60	MACD	C	5	47	0.65	36395.36	0.36
DJIA	max 30	MACD	C	8	18	0.72	35860.82	0.35
HSI	max 60	MACD	C	1	48	0.58	31405.77	0.31
DJIA	benchmark	MACD	C	12	26	0.58	31705.69	0.31
DJIA	max 30	MACD	A	13	26	0.66	25812.65	0.25

Fig 8. Top performing models sorted by Win Rate and Return

As shown in Figure 7, the top performing model resulted in a win rate of 72% and an average return of 35% as indicated by model MACD Group C. These results exceed the project goal of achieving a model that delivers more than 70%-win ratio. Also at this percentage, this model far outperforms any buy-and-hold strategy as it did generate positive returns. It can also be seen that a higher win-rate % does not necessarily mean a higher average return percentage. This was expected, and a higher return percentage is more favorable than a higher win rate percentage, though high values for both was the desired result.

4.2 Results Summary

In summary, as can be seen in Figure 7, the MACD models performed better than the MA models. The MA models were not among the top performing models, and this was especially true when using the DJIA Index. Models also generally performed better when the number of days used for the fast- and slow-moving averages were limited to 30 days as opposed to 60 days. However, models performed better on the HSI Index when the fast- and slow-moving averages were limited to 60 days. The best performing model used an 8-day moving average (fast) and an 18-day moving average (slow) as the parameters. In general, models that used 12 years to train performed better than those that used 3, 6, and 9 years of data. However, among the top models, the best performing models were the ones that used 9 years of data to train. As such, the optimal model was trained on 9 years of data, used the MACD approach, had the fast- and slow-moving averages calculated for 8 and 18 days respectively, and should ideally be used with the DJIA Index.

5 Discussion

Though the objective of this project was met, and the best performing model achieved a consistent 70%-win rate and 35% return percentage in both the DJIA and HSI index, one must wonder this performance, or any of the other models can be generalized to other markets. The results of this project showed that models generally performed better in the DJIA index than in the HSI index. Further research can be done to see how these models will perform in other markets such as NASDAQ, S&P500, or even other smaller or international markets that traders typically engage in. To do this, other methods of technical analysis or machine learning models may have to be employed to further optimize the model.

As such, it would be great to research if artificial intelligence and neural networks can be used to improve the model or even produce an overall better model. Given that the best performing models were the MACD models, constant real time training of the model with artificial intelligence will be worth researching. This is because MACD models assign heavier weights to more recent data, as such a real time training of the model will theoretically keep the model performance optimally high or even improve the performance overtime. With artificial intelligence, this will be possible, and the model can be better trained to optimize the best fast- and slow-moving average parameters to use.

It would also be great to see if the addition of a fault prevention such as “stop loss model” improves the stability of the model. There is a tendency for one to overestimate their ability to engage the market successfully. In some cases, one might purchase an overbought stock or panic and buy and sell prematurely, completely ignoring the buy and sell signals. This is why a fault prevention system built into the model could be beneficial. If such a system was trained using neural network or artificial intelligence, one must wonder how much more effective the model will be as not only will the model give buy and sell signals, but it will adapt to losing streaks to try to minimize them and pass temporary restrictions based on performance patterns.

6 Conclusion

Though challenges to successfully navigating the American stock market or any other stock market will continue to arise, technical analysis will always be a good way to minimize potential pit falls of the stock market. As trends and current events which affect the market continue to change, technical analysis tools and approaches will continue to evolve to account for such changes. Employing the use of machine learning models and technical analysis is a very practical way to establish a pattern of consistent gains without falling prey to the volatility of the stock market and having to put in so much time into monitoring it.

Hyper-tuning model parameters yielded some models that performed well in both the DJIA index and the HSI index. These models trained and backtested to an optimized number of years of data exceeded the goal of a 50%-win rate and yielded a high return percentage on average. Tested against MA models, the MACD models proved to have higher performance peaks, consistency, and market versatility. Though pleased with these results, further research needs to be done to potentially improve the model. Further research can be done to see the usefulness of the model in other markets, the effects of more advanced machine learning techniques, and the effectiveness of adding a fault prevention system to the model.

Overall, all models proved to be more effective than the buy and hold strategy as they take advantage of the volatility in stock prices, prompting the user to buy before prices rise and to sell before they fall. This exponentially increases the profit margins in the long run and minimizes the possibility of experiencing one huge loss if a stock ultimately does not do well. Despite these findings, the risks of engaging the stock market should not be underestimated, and there is no magic formula that guarantees absolute mastery of the market. Keeping up with current trading strategies, and key market influences will always be needed.

Acknowledgments. Special thanks to Dr. Jacquie Cheun-Jensen and Nibhrat Lohia for all their guidance throughout the duration of this project.

References

1. Lam, K. S., Dong, L., & Yu, B. (2019). Value premium and technical analysis: Evidence from the China stock market. Economies, 7(3), 92.
2. Qin, Y., Pan, G., & Bai, M. (2020). Improving market timing of time series momentum in the Chinese stock market. Applied Economics, 52(43), 4711-4725.
3. Lui, Y. H., & Mole, D. (1998). The use of fundamental and technical analyses by foreign exchange dealers: Hong Kong evidence. Journal of International money and Finance, 17(3), 535-545.
4. Li, X., Wu, P., & Wang, W. (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. Information Processing & Management, 57(5), 102212.

5. Cheung, W., Lam, K. S., & Yeung, H. (2011). Intertemporal profitability and the stability of technical analysis: evidences from the Hong Kong stock exchange. Applied Economics, 43(15), 1945-1963.
6. Jiang, F., Tong, G., & Song, G. (2019). Technical analysis profitability without data snooping bias: evidence from Chinese stock market. International Review of Finance, 19(1), 191-206.
7. Hartono, J., & Sulistiawan, D. (2014). The market quality to technical analysis performance: Intercountry analysis. Gadjah Mada International Journal of Business, 16(3), 243-254.
8. Kang, B. K. (2021). Improving MACD technical analysis by optimizing parameters and modifying trading rules: evidence from the Japanese Nikkei 225 futures market. Journal of Risk and Financial Management, 14(1), 37.
9. Lorig, M., Zhou, Z., & Zou, B. (2019). A mathematical analysis of technical analysis. Applied Mathematical Finance, 26(1), 38-68.
10. Milionis, A. E., & Papanagiotou, E. (2011). A test of significance of the predictive power of the moving average trading rule of technical analysis based on sensitivity analysis: application to the NYSE, the Athens Stock Exchange and the Vienna Stock Exchange. Implications for weak-form market efficiency testing. Applied Financial Economics, 21(6), 421-436.
11. Ko, K. C., Lin, S. J., Su, H. J., & Chang, H. H. (2014). Value investing and technical analysis in Taiwan stock market. Pacific-Basin Finance Journal, 26, 14-36.
12. Malkiel, B. G. (1989). Efficient market hypothesis. In Finance (pp. 127-134). Palgrave Macmillan, London.
13. Degutis, A., & Novickytė, L. (2014). The efficient market hypothesis: A critical review of literature and methodology. Ekonomika, 93, 7-23.
14. Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. The journal of Finance, 32(3), 663-682.
15. Fishe, R. P., & Robe, M. A. (2004). The impact of illegal insider trading in dealer and specialist markets: evidence from a natural experiment. Journal of Financial Economics, 71(3), 461-488.
16. Malkiel, B. G. (2005). Reflections on the efficient market hypothesis: 30 years later. Financial review, 40(1), 1-9.
17. Adrian, ran-Moroan. "The relative strength index revisited." African Journal of Business Management 5.14 (2011): 5855-5862.
18. Hung, Nguyen Hoang. "Various moving average convergence divergence trading strategies: A comparison." Investment management and financial innovations 13, Iss. 2 (contin. 2) (2016): 363-369.
19. Martins, Pedro Nuno Veiga. "Technical analysis in the foreign exchange market: the case of the MACD (Moving Average Convergence Divergence) indicator." (2017).
20. Sobreiro, V. A., da Costa, T. R. C. C., Nazário, R. T. F., e Silva, J. L., Moreira, E. A., Lima Filho, M. C., ... & Zambrano, J. C. A. (2016). The profitability of moving average trading rules in BRICS and emerging stock markets. The North American Journal of Economics and Finance, 38, 86-101.