Optimizing the Risk-Return Ratio of a Portfolio using Modern Portfolio Theory and LSTM Forecasting Model for Trading in the Stock Market

NOH, Sang Hyun

Supervised by: Dr. David Rossiter

COMP 4971F - Independent Work

Department of Computer Science and Engineering

The Hong Kong University of Science and Technology

Spring 2022

Abstract

This research will discuss the effectiveness of Modern Portfolio Theory on creating an optimal portfolio with a well balanced risk-return ratio for trading in the stock market, and how the use of forecasting LSTM model could help develop the Modern Portfolio Theory. The traditional Modern Portfolio Theory calculates the expected return and risk using the average and standard deviation of historical price data. The expected return and risk is then used to create a vast number of portfolios forming an efficient frontier when depicted in a scatter plot. The sharpe ratio finds the optimal portfolio that lies on the efficient frontier, and the optimal portfolio can be used to trade. This optimal portfolio will have the optimized risk-return ratio for creating a portfolio. The forecasting LSTM model predicts the future value instead of simply analyzing the historical data like the traditional way, so the implementation of the LSTM deep learning model can further develop the effect of the Modern Portfolio Theory. Throughout the research, the ideal number of individual securities to create a portfolio using Modern Portfolio Theory, balance of train and test data to predict stock prices using forecasting LSTM model, effect of using Modern Portfolio Theory to rebalance existing market portfolios like Dow Jones Industrial Average are also investigated. As a result, the Modern Portfolio Theory was effective in creating a portfolio that can perform better than Dow Jones 30 or other diversified portfolios in the industry. The CAGR of the optimal portfolio was 14.03% from 2017 to 2021 while the CAGR of Dow 30 was 11.70%. Also, the result was the best for a portfolio with around 10 individual securities, while it shows similar results for a portfolio with around 5 individual securities. The application of forecasting LSTM model drastically improved the performance of Modern Portfolio Theory, as its CAGR recorded 87.78% while the equally balanced portfolio only had a CAGR of 26.58%.

Table of Contents

Table of Figures

1. Introduction

1.1 Stock Market and Diversification

The stock market is always volatile and is regarded as a high-risk, high-return market. Especially, investing in a single stock holds a great risk. This risk can be reduced by making a diversified portfolio consisting of a variety of assets. It limits the exposure by holding a single asset. In this paper, we focus on diversification within the stock market.

1.1.1 Return and Risk of Individual Securities

Expected return of individual securities is a return that an individual expects a stock to earn over the next period. It can be estimated by the average return per period a security has earned in the past.

$$Expected\ Return\ =\ Average\ return\ per\ period$$

The risk of individual securities can be measured by the variance or standard deviation, which shows the volatility of a security. It considers risk as how much the security's actual return has varied from the security's expected return.

$$Risk\ =\ Var\ =\ (Security's\ return\ -\ Expected\ Return)^2$$

1.1.2 Return and Risk of a Portfolio

Expected return of a portfolio is the weighted average of expected return of individual securities.

$$E(R_p)\ = X_A \cdot E(R_A)\ + X_B \cdot E(R_B)$$

E(R) : Expected Return,

X_k : Weight of asset k, $E(R_k)$: Expected Return of asset k

The risk of a portfolio depends on both variance of individual securities and covariance between them. The positive relationship between individual securities increases variance of a portfolio, and vice versa.

$$Risk(Portfolio) = X_A^2 \sigma_A^2 + 2 \cdot X_A X_B \sigma_{A,B} + X_B^2 \sigma_B^2$$

X_k : Weight of asset k, σ_k : Standard Deviation of asset k, $\sigma_{A,B}$: Covariance of asset A and B

1.1.3 Diversification Effect

The risk in a stock market can be categorized into a systematic risk and an unsystematic risk. Systematic risk is the market risk that affects a large number of assets. Unsystematic risk only affects a single asset and it is unique to individual companies. Diversification reduces the unsystematic risk and eventually reduces the risk of a portfolio. As long as two securities are not perfectly correlated, the risk of the portfolio is less than the weighted average of the risk of the individual securities. Despite this may lead to lower return, the mitigated risk assures safer investment.

1.1.4 Dow 30

Dow Jones Industrial Average, or Dow 30 is a common U.S. stock market benchmark composed of 30 large, publicly traded companies in the New York Stock Exchange (NYSE). It is an example of a diversified portfolio which holds top 30 individual stocks in NYSE.

1.2 Modern Portfolio Theory

Modern Portfolio Theory refers to a method for assembling an asset portfolio to maximize the expected return for an acceptable level of risk. The theory was devised by Harry Max Markowitz in 1952 and was awarded a Nobel Prize for the work. The core component of Modern Portfolio theory is diversification, and the goal is to find the optimal mix of individual assets by using the relationship between expected return and risk.

1.2.1 Efficient Frontier

Efficient Frontier is the key idea in Modern Portfolio Theory. It graphically shows portfolios that maximize the return for a specific risk that the investors can take. Return of a portfolio is dependent on the combinations that form a portfolio.

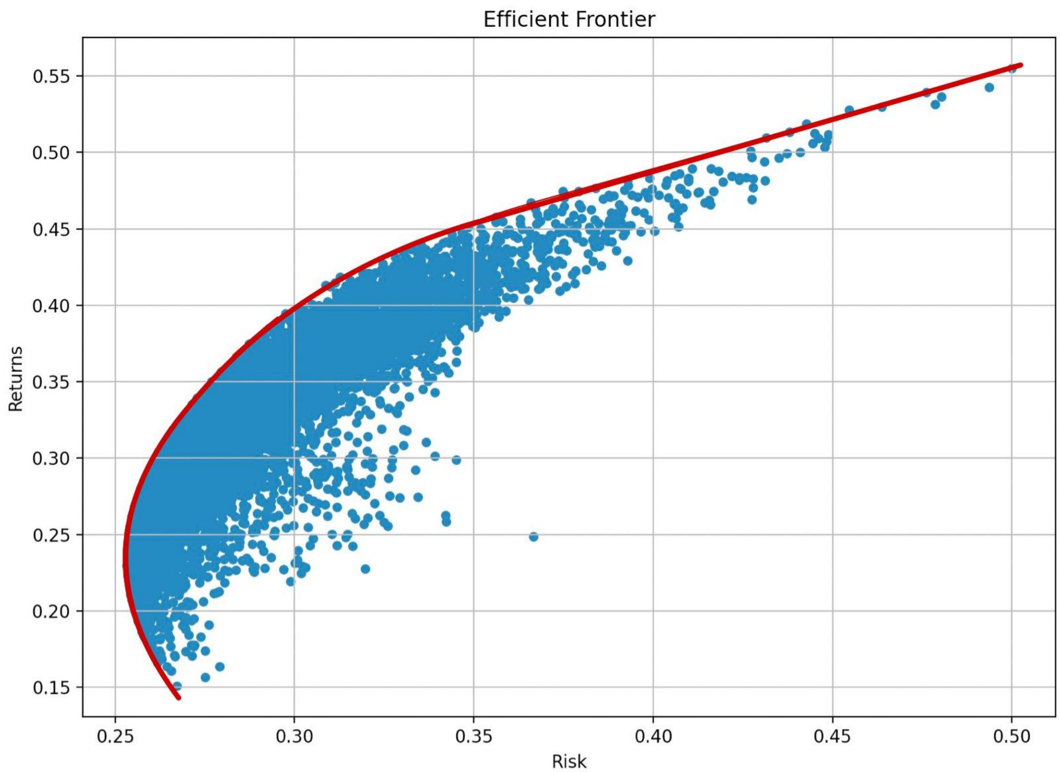

Figure 1. *Efficient frontier drawn in a red line*

In Figure 1, each blue dot represents a portfolio, and they all have different weights for the components of a portfolio. The red line represents the efficient frontier, and the portfolios that lie on the efficient frontier are expected to give the best return for a specific risk. However, any other portfolios that lie within the efficient frontier isn't ideal as they take the same risk when higher expected return is possible.

1.2.2 Sharpe Ratio

There are numerous portfolios that lie on the efficient frontier, and the best portfolio should be selected among them in order to choose the weights for trading in the stock market. Sharpe ratio is used to compare between these portfolios as it calculates the expected return per risk.

$$Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p}$$

R_p : Return of a Portfolio, R_f : Risk-free rate, σ_p : Risk of a Portfolio

In this paper, the risk-free rate is assumed to be 0 for less complex calculation, and sharpe ratio is calculated by dividing the expected return of a portfolio by the risk of a portfolio. For example, if the expected return is 7% and the risk of a portfolio is 5%, then the sharpe ratio is 1.4. The higher the sharpe ratio, the higher the return for a given risk.

1.2.3 Portfolio Optimization

The optimum portfolio and its weight for each individual stock is the portfolio with the highest sharpe ratio that lies on the efficient frontier. In Figure 2, the portfolio with the optimum portfolio is '★' and the portfolio with lowest risk is '✖'.

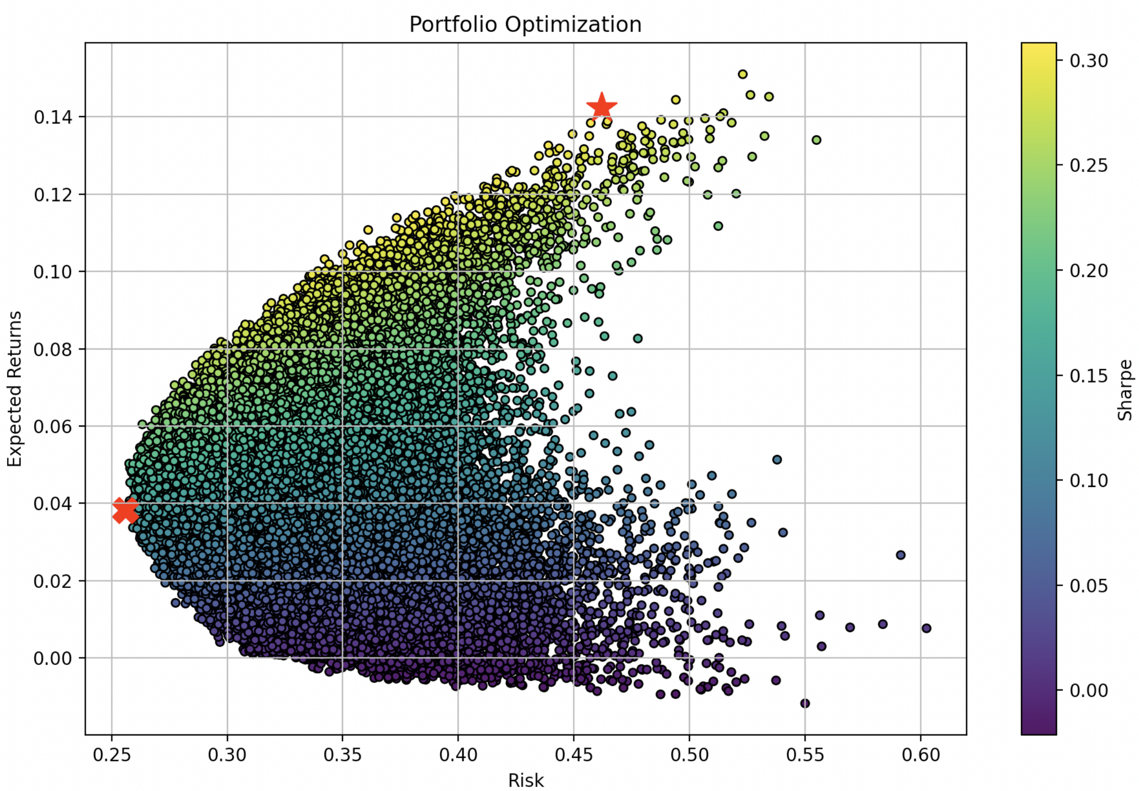

Figure 2. *Visualization of the portfolio optimization*

1.3 LSTM Forecasting Model

Long-Short Term Memory network, or LSTM in short, is used in the fields of deep learning. It can be used for time-series forecasting to predict data with time dimensions like stock prices.

1.3.1 Time-Series Forecasting

Time-series forecasting is predicting the future value using the previously observed values. It is widely used for non-stationary data, which is the data that has varying mean and standard deviation.

1.3.2 Recurrent Neural Network (RNN)

A neural network is an algorithm used to recognize the underlying relationships in a set of data by mimicking the operations done in the human brain. Recurrent Neural Network, or RNN, is a class of neural networks. A node in the RNN performs a calculation and returns an output value, and the output is used in the next node as an input. This is recurring as it continues to reuse the output from a previous step as an input of a next step.

1.3.3 Long-Short Term Memory (LSTM)

LSTM is a special kind of RNN that has a recurrent node and an internal state. The internal state is used as a working memory space in order to store and retrieve information over many time stamps. The calculation in each node of the LSTM network involves the input value and the internal state. The calculation is not only used to return a new output value, but it also updates the internal state. LSTM has a special parameter called gates, which is used to control how much of internal state information is used for calculation in each node. The idea is to keep some meaningful data and to forget useless data. The drawback of LSTM is that it tends to be slow for training, because it uses a lot of parameters.

1.4 CAGR

Compound Annual Growth Rate (CAGR) is the ratio that provides a constant rate of return over the time period. It is one of the most accurate ways to determine returns that can vary over time. High CAGR indicates high return from the investment.

$$CAGR = \left(\left(\frac{E_v}{B_v}\right)^{\frac{1}{n}} - 1\right) \times 100$$

B_v : Beginning Value, E_v : Ending Value, n: number of years

1.5 Maximum Drawdown

Maximum Drawdown is the maximum observed loss before a new peak is attained. It indicates the downside risk over a time period.

$$Maximum\ Drawdown\ =\ \frac{V_T - V_P}{V_P}$$

V_P : Peak Value, V_T : Trough Value

1.6 Objective

This research project aims to test if modern portfolio theory can effectively find the optimum portfolio that can generate consistent return. It also tests if the application of forecasting LSTM model in the modern portfolio theory is effective. The effectiveness is assessed using CAGR and drawdown.

2. Methodology

2.1 Data Collection

In order to determine the effectiveness of the Modern Portfolio Theory and the LSTM forecasting model, real data from the stock market was collected. Because this paper focuses on creating the optimal portfolio, historical price data of many individual securities and some market portfolios like Dow 30 were collected.

Table 1. *Data of Dow 30 from 2000-01-03*

Date	High	Low	Open	Close	Volume	Adj Close
2000-01-03	11522.009765625	11305.6904296875	11501.849609375	11357.509765625	169750000	11357.509765625
2000-01-04	11350.0595703125	10986.4501953125	11349.75	10997.9296875	178420000	10997.9296875
2000-01-05	11215.099609375	10938.669921875	10989.3701171875	11122.650390625	203190000	11122.650390625
2000-01-06	11313.4501953125	11098.4501953125	11113.3701171875	11253.259765625	176550000	11253.259765625
2000-01-07	11528.1396484375	11239.919921875	11247.0595703125	11522.5595703125	184900000	11522.5595703125
2000-01-10	11638.2802734375	11532.48046875	11532.48046875	11572.2001953125	168180000	11572.2001953125
2000-01-11	11663.099609375	11502.7001953125	11568.4697265625	11511.080078125	177300000	11511.080078125

The data was acquired from Yahoo Finance in US dollars as a dataframe in python, and it was converted to a csv file as shown in Table 1. It has information of date, opening price, lowest price, highest price, closing price, adjusted closing price, and the volume of transaction. Daily close prices were used for the making of Portfolio. The daily close price from 2000 of the Dow 30 is shown in Figure 3.

Figure 3. *Graph of collected data vs Graph of real data for comparison (Dow 30)*

2.2 Investment Method

The key idea of this paper is to create a diversified portfolio with the optimal balance of weight for each individual security. New portfolio will be created every month using the past three years of closing price data of individual securities. The reason why it is using past 'three' years of data is for the better performance of forecasting LSTM model, which will be explained in section 2.5. Expected return of individual securities will be calculated using either the forecasting LSTM model or average of past data, while the risk of individual securities will be simply calculated by the variance of past data. The optimal balance of weight for a diversified portfolio will be found using the expected return and risk, and the investment will be made for each individual security according to its weight. Each security will be held for a month until a new optimal portfolio is created. The cost of buying or selling is ignored, and the change in closing price every month is considered as the profit or loss. After investing for a few months, the optimal portfolio will be compared with an equally balanced portfolio by the CAGR to measure its performance.

2.3 Data Preprocessing

The closing price of individual securities that forms a portfolio needs to be collected into one dataframe. Because three years worth of data is required to invest for the following month, with 8 years of data, investments can be tested for 5 years. The closing prices of the first 5 days for four different companies listed in Dow 30 are shown in Table 2.

Table 2. *Closing price of Tesla, P&G, J.P. Morgan, and U.S. Real Estate ETF from 2014*

Date	TSLA	PG	JPM	IYR
2014-01-02	30.020000457763672	80.54000091552734	58.209999084472656	62.97999954223633
2014-01-03	29.91200065612793	80.44999694824219	58.65999984741211	63.34999847412109
2014-01-06	29.399999618530277	80.63999938964844	59.0	63.61000061035156
2014-01-07	29.871999740600582	81.41999816894531	58.31999969482422	63.83000183105469
2014-01-08	30.256000518798828	80.23999786376953	58.869998931884766	63.72999954223633

After collecting the data for 8 years of data, two more specific data frames need to be created iteratively. The first data frame will have the first three years of data, which is used to calculate expected return and risk. The second data frame will have one month of data, which is used to test the performance of the investment made. In terms of LSTM, the first data frame will be the train data and the second data frame will be the test data. So everytime it creates a new portfolio, the train data and testing data are required.

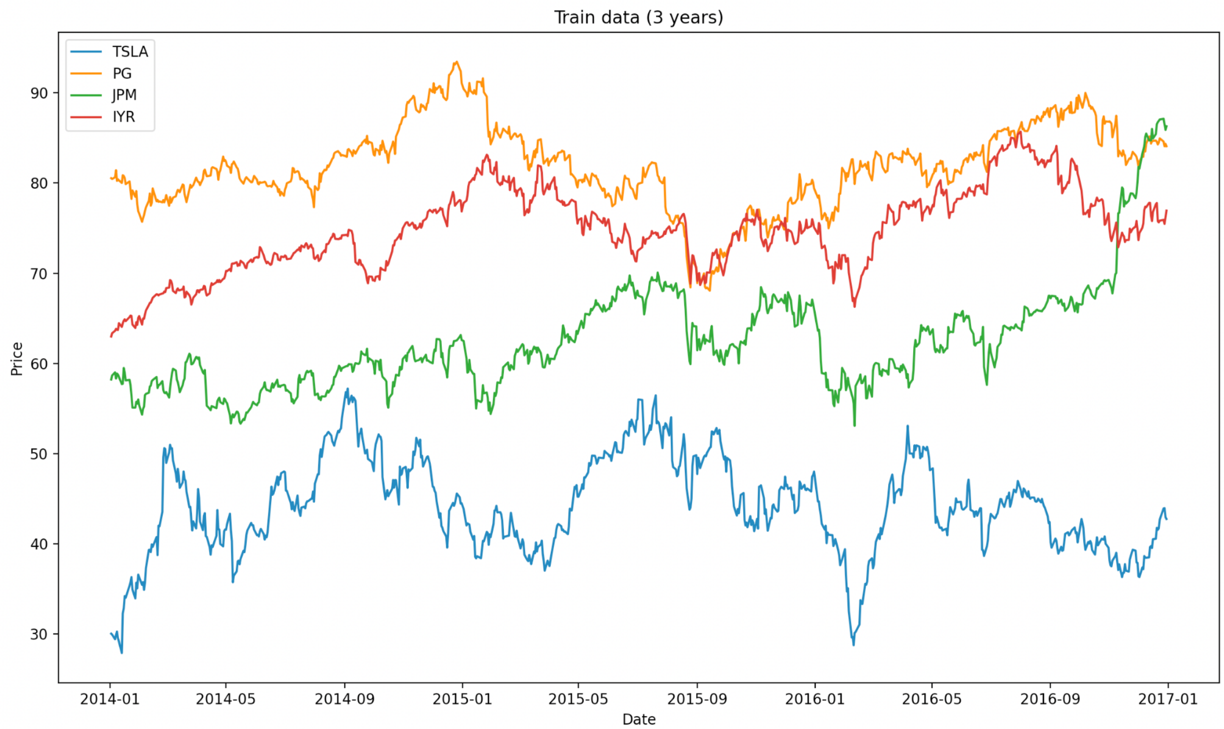

Figure 4. *Train data of 4 different individual securities*

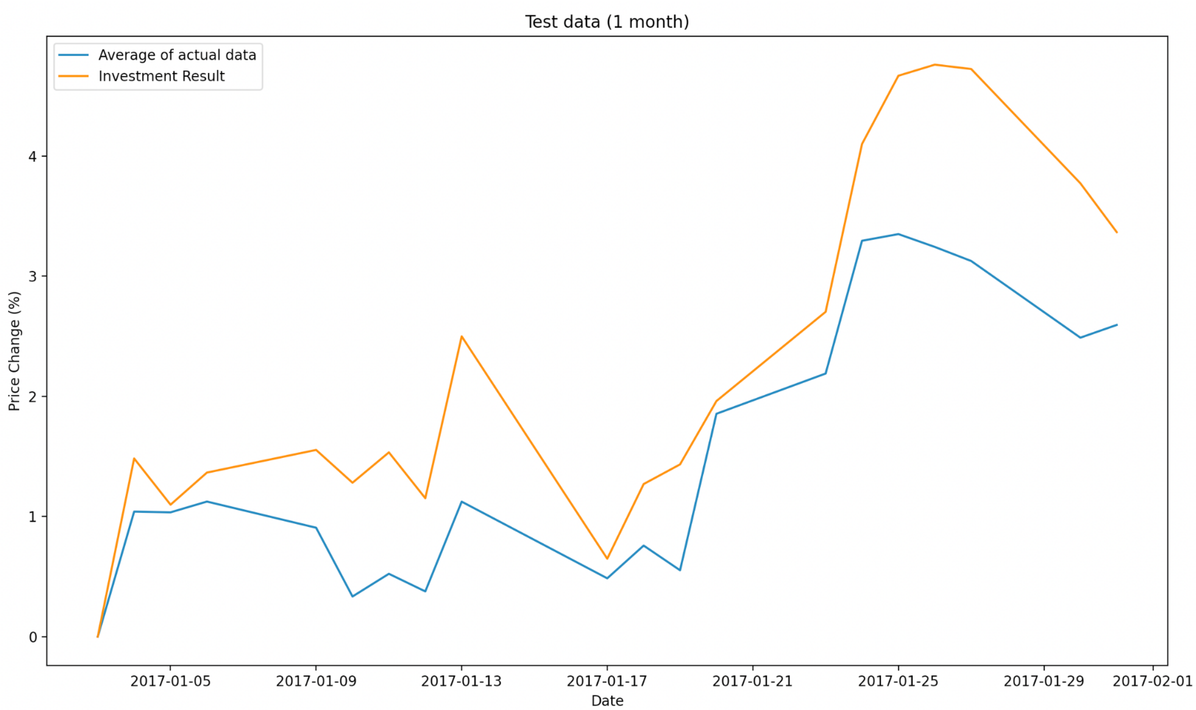

Figure 5. *Equally weighted test data used to test performance of investment*

When using the second data frame to test the performance of the investment made, the closing price of each stock is weighted equally and the average price was calculated to be used as a comparison standard. The closing price for the first three years of each individual securities from Table 4 are shown in Figure 4. Figure 4 corresponds to the first data frame which is used as the train data. Figure 5 corresponds to the second data frame which is used as the test data. The investment is made according to the optimal balance of each securities calculated from the Modern Portfolio Theory, and the price changes from investment are shown in the yellow line, while the equally weighted price changes are shown in the blue line.

2.4 Calculating Expected Return and Risk from Past Data

Before implementing the forecasting LSTM model to predict the return of each individual stock, the expected return was calculated using the past data because it can be calculated a lot faster. With the past three years of data, the daily return was found by the average percentage change, and the monthly return was found by multiplying the daily return by 21. Each month is considered to have 21 trading days because there are an average of 252 trading days in a year, and if it is divided by 12 months, it is exactly 21. Then, the daily risk was found by the daily covariance of closing price, and the monthly risk was found by multiplying the daily risk by 21. The expected return and risk for the first month using the data for the previous 3 years are shown in Figure 6.

MONTHLY	RETURN
TSLA	0.012493
PG	0.005630
JPM	0.018720
IYR	0.005642

MONTHLY RISK

	TSLA	PG	JPM	IYR
TSLA	0.014085	0.000668	0.001479	0.001133
PG	0.000668	0.001665	0.000633	0.000703
JPM	0.001479	0.000633	0.003781	0.000669
IYR	0.001133	0.000703	0.000669	0.001589

Figure 6. *Monthly Return and Risk calculated from past data, where a value of 0.01 means 1%*

2.6 Forecasting LSTM Model to predict return

The forecasting LSTM model can be used to predict the return in a coming month, and this data can be used in the Modern Portfolio Theory. Because it is predicting the future data instead of simply finding the average of past data, the expected return calculated by the forecasting LSTM model should help the investment perform better.

A multi-layer LSTM recurrent neural network is used to predict the closing price for the future 21 days which can be considered as 1 month. Tesla's stock price was used in this example for demonstration.

The first step is to split the data into training sets and testing sets to avoid any overfitting problem. This is done as explained in section 2.3. The balancing of train and test data is important because a lack of train data can lead to poor prediction, while having a lot of train data will cause the computing time to be too long. In figure 7, it shows the performance of the forecasting LSTM model with 1 year, 2 years, 3 years, and 4 years of close price of Tesla as train data to predict the closing price of 1 month ahead. It can be seen that the train data with 1 year has poor performance while the train data with 4 years seems to be overfit. Also, the LSTM forecasting model takes a lot of time, so it is better to use less data for train data. Therefore, it is reasonable to use the 3 years of data to predict the price of 1 month ahead.

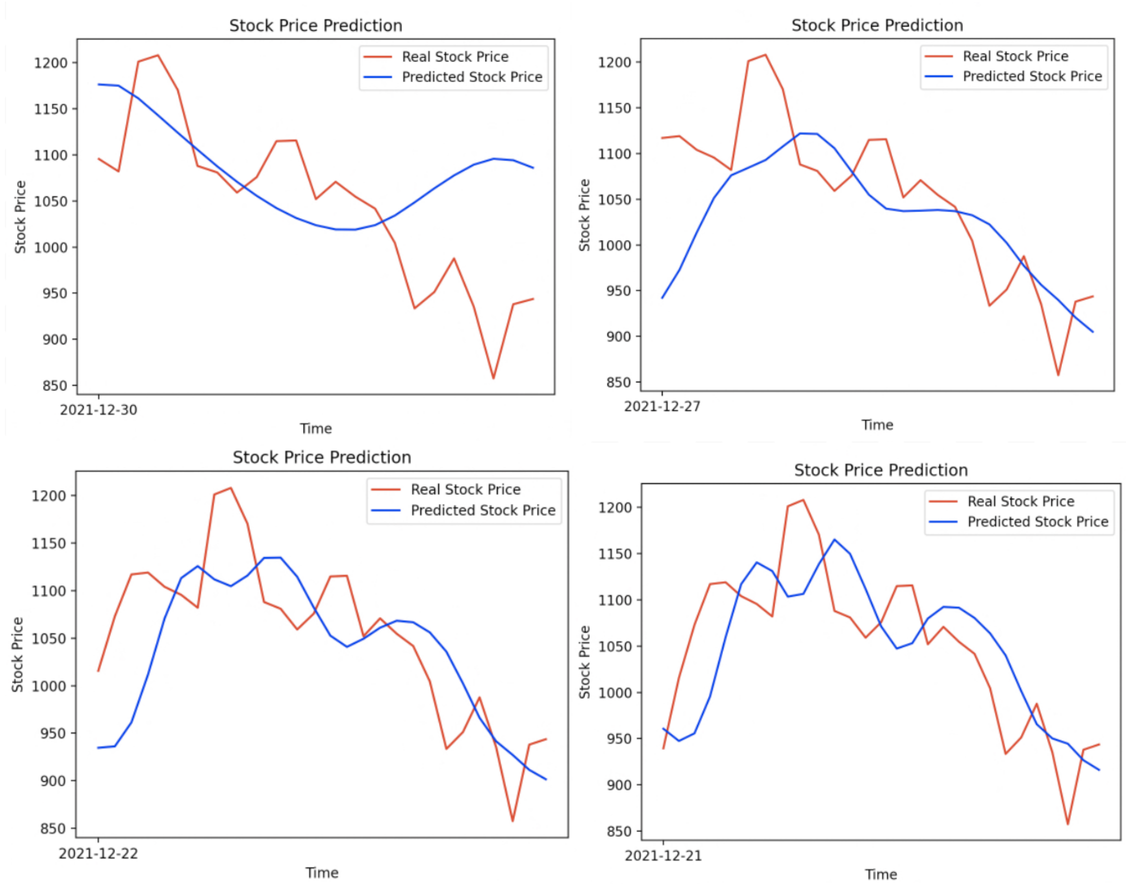

Figure 7. *Predicted stock price of Tesla from LSTM using 1, 2, 3, 4 years of data shown in top left, top right, bottom left to bottom right.*

2.7 Monte Carlo Simulation

Using the expected return and risk calculated, we can create a randomized portfolio that has random weights for each individual security. We made 100,000 portfolios using the Monte Carlo Simulation, which is a model used to predict the probability of different outcomes. These 100,000 portfolios will then create an efficient frontier as explained by the Modern Portfolio Theory, and it would be ideal to make a portfolio that lies on the efficient frontier. If there is a specific risk value that an investor is willing to take, then they can choose the portfolio that lies on the efficient portfolio for the corresponding risk. If an investor wants to invest in a safe portfolio, they can choose the portfolio with the lowest risk. This portfolio is marked with '✖' in Figure 8. The most efficient portfolio is the portfolio with the largest sharpe ratio, which is marked with '★' in Figure 8. In this paper, the most efficient portfolio is selected to use for trading in the stock market.

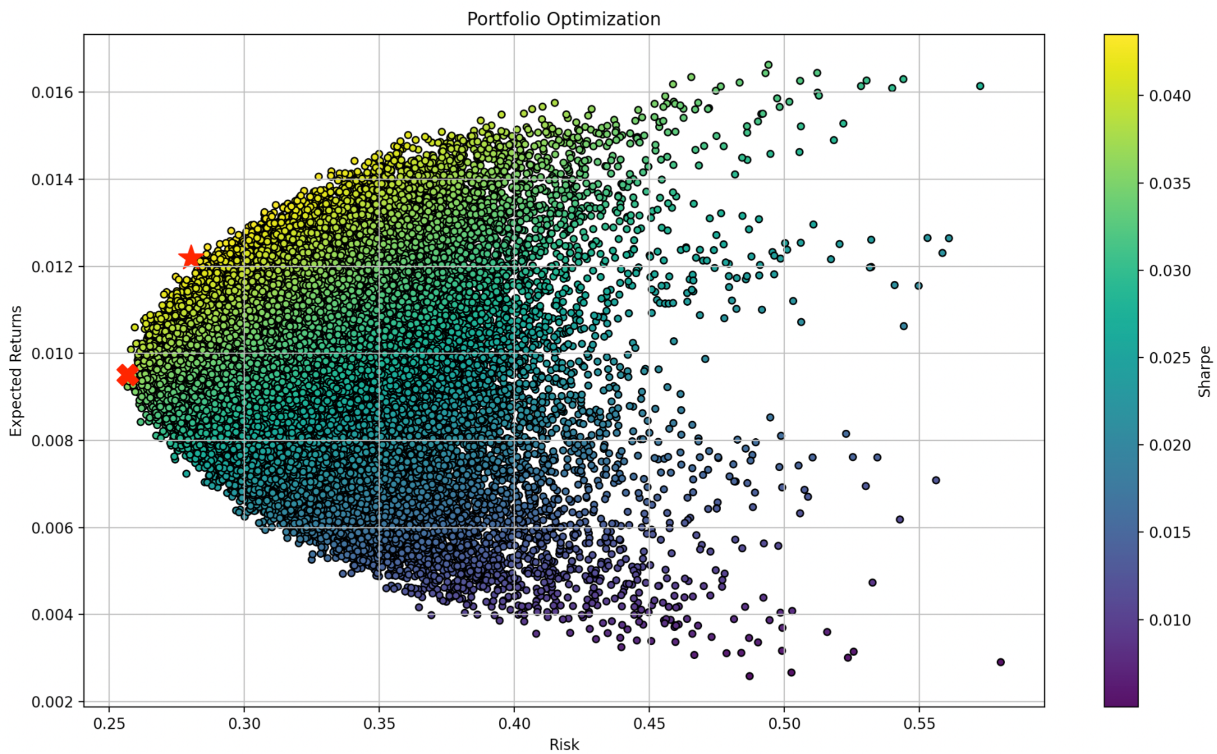

Figure 8. *100,000 portfolios created with Monte Carlo Simulation*

3. Data Analysis

3.1 Finding Number of Individual Securities to use in a Portfolio

Using the method above, it is now possible to find an optimal portfolio with any type of individual securities selected as long as historic price data is given to calculate the expected return and risk. But we need to identify how many individual securities we need to use in a portfolio. This can be tested by creating an efficient frontier using Monte Carlo simulation, and seeing how the number of individual securities in a portfolio changes the shape of an efficient frontier.

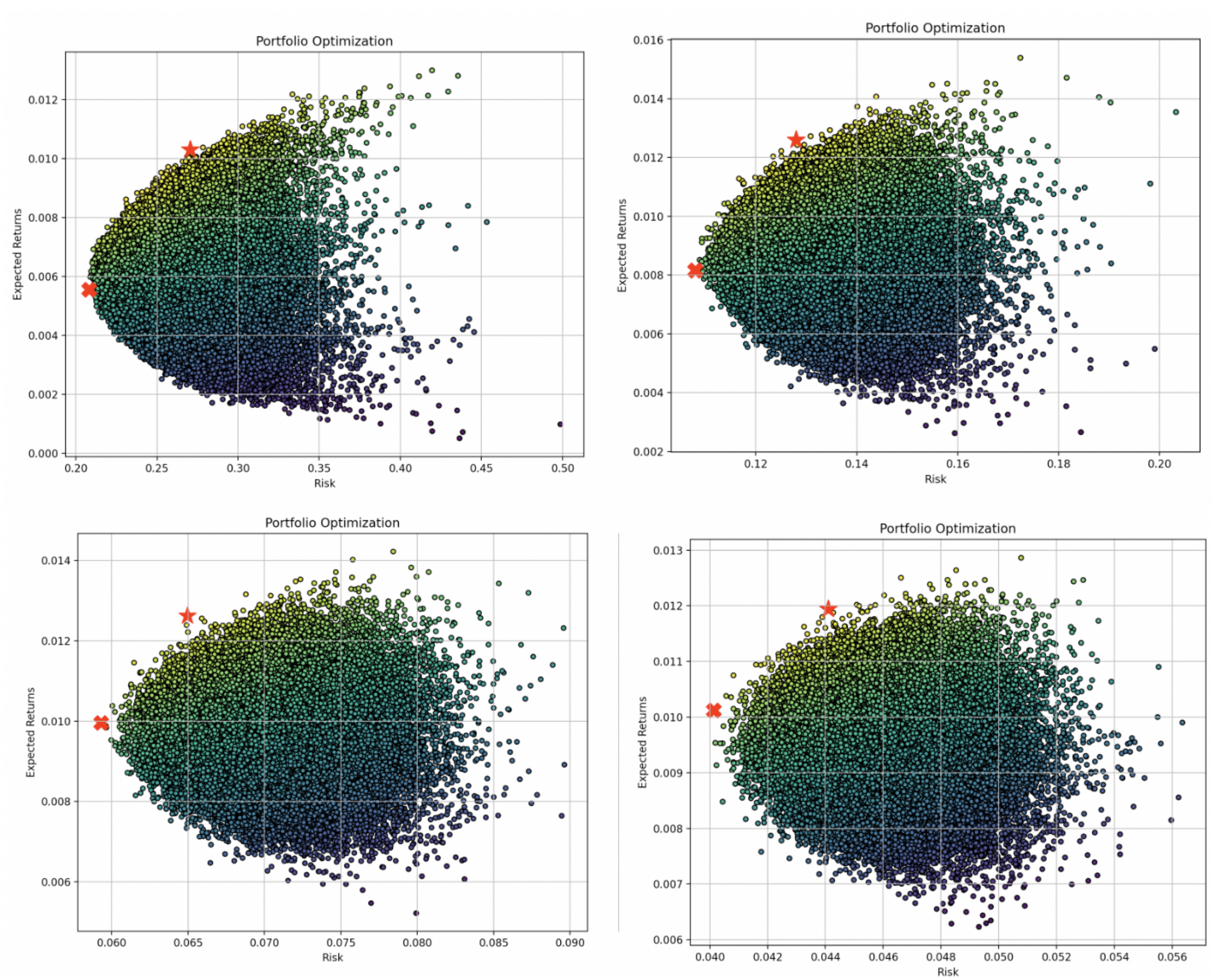

Figure 9. *100,000 portfolios created for 5, 10, 20, and 30 individual securities shown in top left, top right, button left to bottom right.*

It can be seen in Figure 9 that the less the number of individual securities, the less anomalies there are near the efficient frontier. These portfolios were created without the LSTM application, simply depending on the original Modern Portfolio Theory using past data to find expected return. The portfolio with 5 individual securities has a clean curvature of efficient frontier shape while the portfolio with 30 individual securities has a lot of random portfolios around the efficient frontier. So it can be judged that the Modern Portfolio Theory can be effectively applied for creating portfolios with around 5 individual securities.

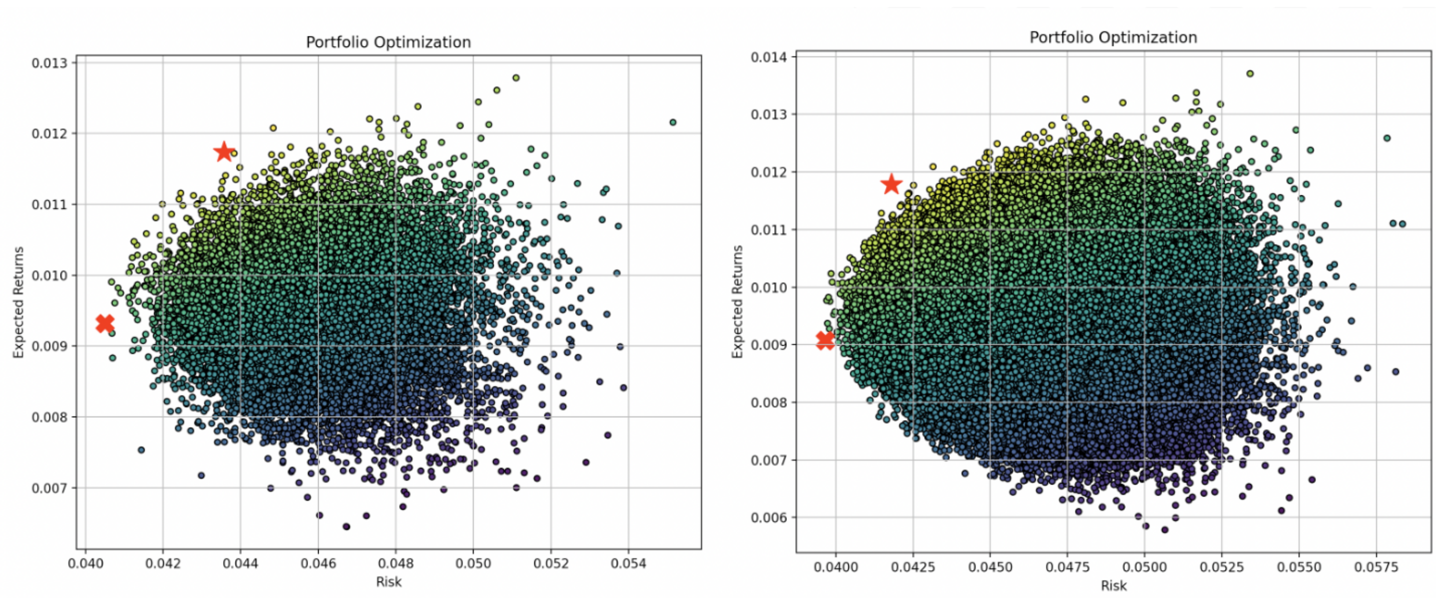

Figure 10. *20,000 portfolios and 10,000,000 portfolios created using 30 individual securities*

The reason why the Modern Portfolio Theory performs better with less number of individual securities is due to the number of portfolios formed during the Monte Carlo Simulation. This can be visualized in Figure 10, showing the effect of decreasing the number of portfolios to 20,000 or increasing the number of portfolios to 10,000,000. This shows that the Modern Portfolio Theory can still be applied if we simply increased the number of portfolios created in the Monte Carlo Simulation, but this takes a lot of computing time.

3.2 Application to Dow 30

Before testing the Modern Portfolio Theory with around 5 individual securities, it can be tested with 30 individual securities by comparing with Dow 30. This is to test the performance for investing in the top performing companies that are listed in the NYSE. It can also give a general idea of how it performs relative to the market index. Dow 30 was chosen here instead of some broader market index like S&P 500 because a portfolio is said to perform best when it has around 20 to 30 securities. Dow 30 constitutes 30 prominent companies listed in the NYSE, and it is a price-weighted stock index. Whole list of components of Dow 30 and its historic price can be acquired, so the Modern Portfolio Theory was tested with this data. 8 years of data for every component of Dow 30 were collected from 2014 to 2021, to test the Modern Portfolio Theory for 5 years from 2017 to 2021. The stock market experienced a huge drop due to covid-19 during this period while having some constant growth in other periods of time. So the 5 years would be a good time to test for its CAGR and maximum drawdown.

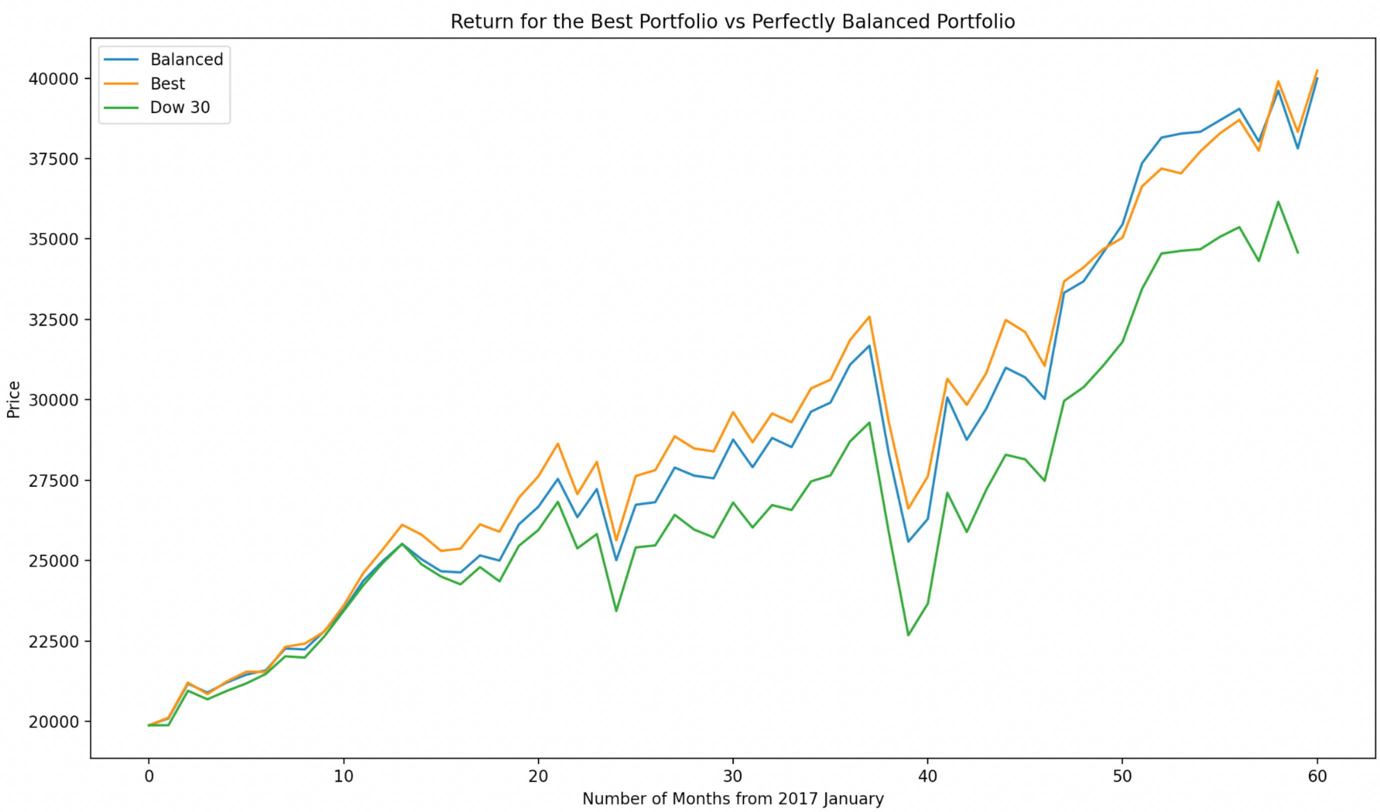

Figure 11. *Return of optimal portfolio, equally weighted portfolio, and price of Dow 30*

Table 3. CAGR and Maximum drawdown of different portfolios - Dow 30

	Total Profit	CAGR	Maximum Drawdown
Optimal Portfolio	**92.78%**	**14.03%**	**- 10.10%**
Equally-weighted Portfolio	90.19%	13.72%	- 10.63%
Dow 30	73.93%	11.70%	- 12.31%

According to the data from Figure 11 and Table 3, the optimal portfolio created using the Modern Portfolio Theory performed the best by making the largest total profit and CAGR compared to real Dow 30 and equally-weighted portfolio. The maximum drawdown during the covid-19 outbreak was also the smallest at 10.10%. The difference between the optimal portfolio and the equally-weighted portfolio isn't substantial, but it can be seen that the Modern Portfolio Theory is still effective despite forming imperfect efficient frontier as shown in section 3.1.

3.3 Application to 11 sectors in NYSE

The analysis of the effect of Modern Portfolio Theory can be tested with the reduced size of portfolio constituents of 11. The Global Industry Classification Standard (GICS) is created by Standard and Poor's (S&P) to sort publicly listed companies in NYSE into 11 sectors based on their business activity. Instead of investing in the ETF which is already a diversified portfolio, we can select the biggest holdings of each sector and invest in them to see the effect. The 11 sectors from GICS and the biggest weighted company for each sector is shown in Table 4.

Table 4. 11 Sectors of GICS and its biggest weighted company

Sector	Ticker	Company	Ticker	Net Asset
Information Technology	XLK	Apple Inc.	AAPL	23.45%
Health Care	XLV	UnitedHealth Group Inc.	UNH	9.20%
Financials	XLF	JPMorgan Chase & Co.	JPM	9.45%
Consumer Discretionary	XLY	Amazon.com Inc.	AMZN	23.17%

Communication Services	XLC	Meta Platforms Inc.	FB	19.88%
Industrials	XLI	Union Pacific Corp.	UNP	5.77%
Consumer Staples	XLP	Procter & Gamble Co.	PG	15.68%
Energy	XLE	Exxon Mobil Corp.	XOM	22.23%
Utilities	XLU	NextEra Energy Inc.	NEE	15.81%
Real Estate	XLRE	Prologis Inc.	PLD	11.44%
Materials	XLB	Linde PLC	LIN	16.31%

The portfolios created with 11 individual stocks using the Modern Portfolio Theory can be compared with the Standard and Poor's 500 market index which has 500 large companies in NYSE. Although it has a different number of individual stocks in a portfolio, it would be a good comparison as they are the same company which divided the industry into 11 sectors.

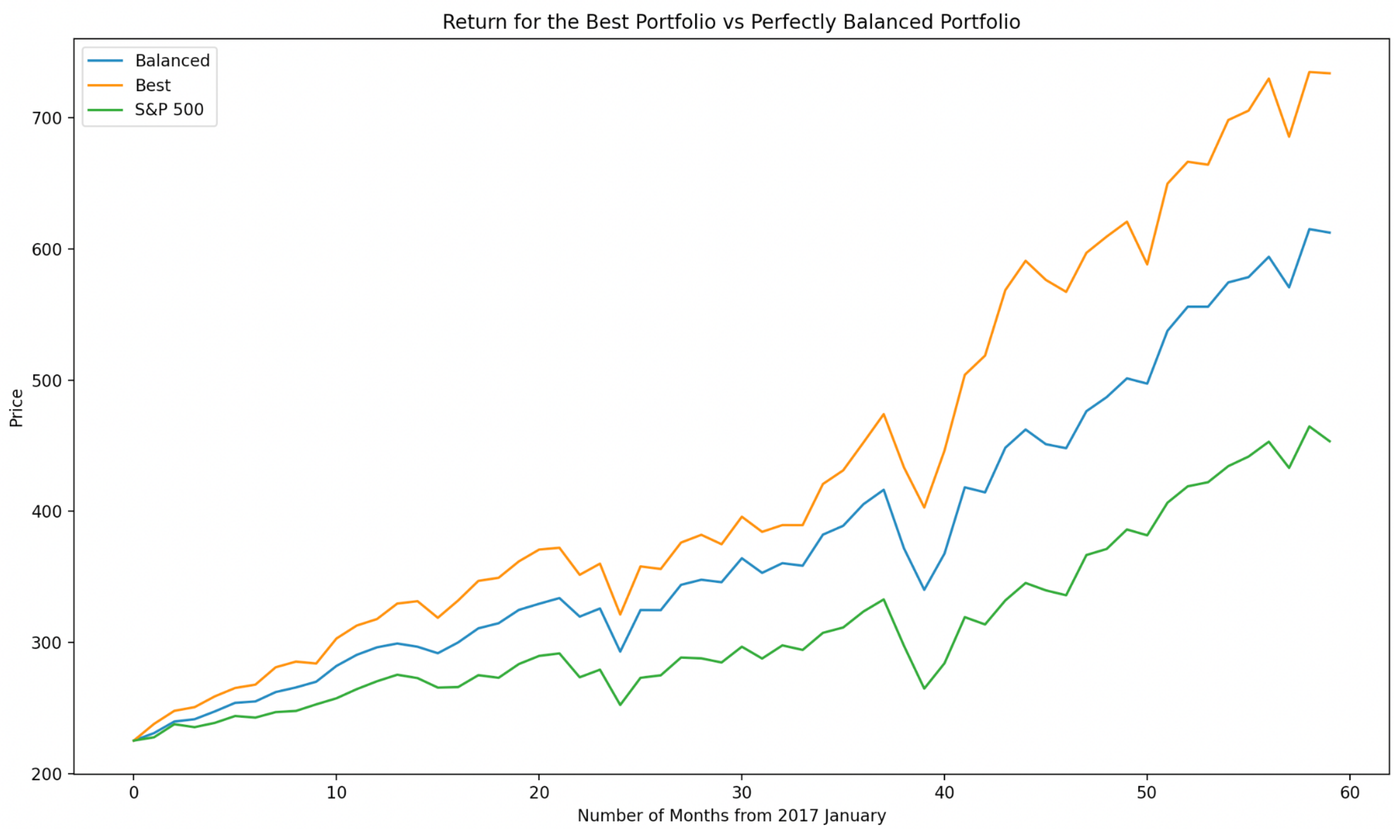

Figure 12. *Return of optimal portfolio, equally weighted portfolio, and price of S&P 500*

Table 5. *CAGR and Maximum drawdown of different portfolios - 11 sectors*

	Total Profit	CAGR	Maximum Drawdown
Optimal Portfolio	**225.85%**	**26.65%**	**- 8.58%**
Equally-weighted Portfolio	171.92%	22.15%	- 10.73%
S&P 500	101.31%	15.02%	- 10.96%

According to the data from Figure 12 and Table 5, the optimal portfolio created using the Modern Portfolio Theory performed the best by making the largest total profit and CAGR compared to real Dow 30 and equally-weighted portfolio. The CAGR was around twice the S&P 500's stock data and this shows how the less number of individual securities leads to the better performance of Modern Portfolio Theory. The maximum drawdown was also less by 2% at -8.58% while the equally-weighted portfolio and S&P 500 were around -11%.

3.4 Application to 4 sectors in NYSE - AMZN, PG, JPM, PLD

The number of individual securities in a portfolio was reduced further to see the effect. Four sectors from the 11 sectors of GICS were selected to make a portfolio. They are Amazon, P&G, J.P. Morgan and Prologis which correspond to Consumer Discretionary, Consumer Staples, Financials, and Real Estate. The optimal weight of the four individual securities for the first 7 months are shown in Table 6. The weights from each company in a month adds up to 1.

Table 6. *Optimal Weight of Individual Securities for 4 sectors in NYSE*

	AMZN	PG	JPM	PLD
0	0.4283641144378387	0.11843021616835746	0.2799069543726888	0.17329871502111507
1	0.5503543591832104	0.03539001280538533	0.21608583636510875	0.19816979164629553
2	0.40565849718642233	0.1550499608274371	0.25509813853928903	0.18419340344685162
3	0.5458061445791264	0.11101668182377719	0.18645853145199798	0.15671864214509842
4	0.5473112187129981	0.04197383628694283	0.3086164974512836	0.10209844754877544
5	0.5329690644719637	0.0948430102223853	0.1717379662890625	0.20044995901658863
6	0.4564286905571476	0.063006711769660 1	0.24426102975857675	0.23630356791461543

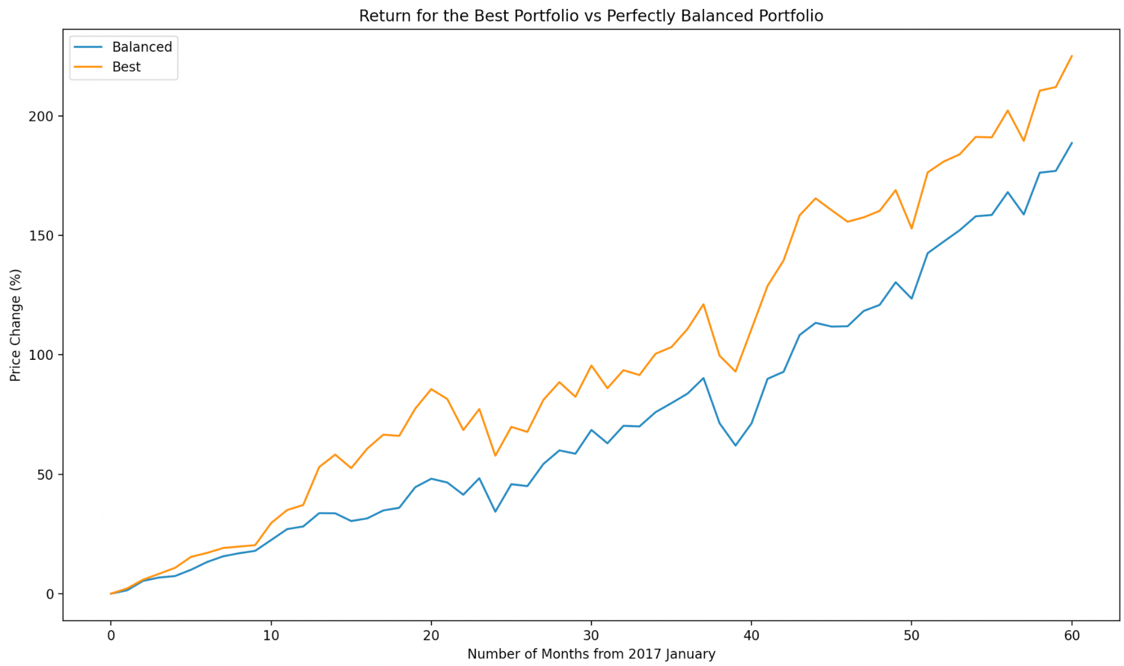

Figure 13. *Comparison of optimal portfolio and equally-weighted portfolio for 4 sectors*

Table 7. *CAGR and Maximum drawdown of different portfolios - 4 sectors*

	Total Profit	CAGR	Maximum Drawdown
Optimal Portfolio	**225.06%**	**26.59%**	**- 9.73%**
Equally-weighted Portfolio	188.64%	23.62%	- 9.92%

The result from Figure 13 and Table 7 shows that the portfolio composed by 4 individual stocks perform similarly as a portfolio composed by 11 individual stocks. This is because there isn't a great difference in the shape of the efficient frontier shown in Figure 9. However, the maximum drawdown was better for 11 individual stocks which should be due to the lack of diversification effect to reduce the risk of the portfolio. From the actual application of the investment method, it can be concluded that the Modern Portfolio Theory is most suitable for a portfolio with around 10 individual securities, while it shows similar results for a portfolio with around 5 individual securities.

3.5 Effect of Forecasting LSTM model

The use of the forecasting LSTM model can be tested by replacing the expected return calculated by the average monthly return into the expected return predicted by the LSTM machine learning model. The effect of the forecasting LSTM model is tested on the portfolio with 4 individual stocks identical to the ones from section 3.4. A small amount of individual stocks is preferred due to its extremely long computational time to predict the price using the LSTM model.

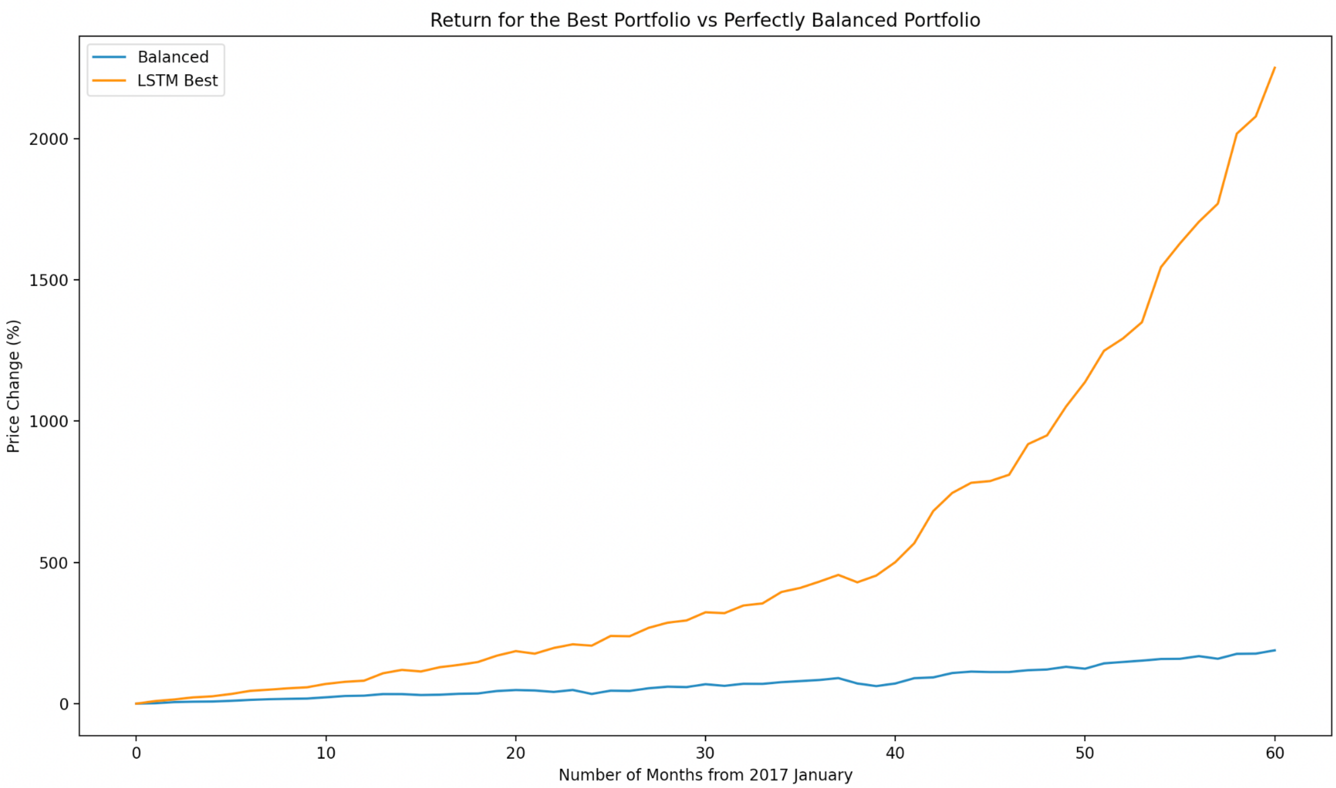

Figure 14. *Comparison of optimal portfolio created by forecasting LSTM for 4 sectors*

Table 8. *CAGR and Maximum drawdown of different portfolios - LSTM prediction*

	Total Profit	CAGR	Maximum Drawdown
Optimal Portfolio	**2250.68%**	**88.03%**	**- 4.71%**
Equally-weighted Portfolio	188.64%	23.62%	- 9.92%

The effect of the forecasting LSTM model for predicting expected return in Modern Portfolio Theory is shown in Figure 14 and Table 8. The CAGR from the optimal portfolio using LSTM prediction was 88.03% which was more than three times the equally-weighted portfolio. The maximum drawdown was also almost half from the equally-weighted portfolio. This could be achieved by the high performance of the LSTM model to correctly predict the future value in a month. The expected return and the balance for each individual security according to the expected return for the first 5 months is shown in Table 9 and 10. It shows that if the stock price is predicted to fall, the weight for that stock would be minimal to reduce the loss from making investment.

Table 9. *Expected Return from the forecasting LSTM model*

	AMZN	PG	JPM	PLD
0	0.071323186	0.024130415	-0.00353425	-0.050689287
1	0.017454257	0.044525772	0.08118538	0.036373843
2	0.007371442	-0.005183988	-0.03304861	-0.002974625
3	0.048040237	-0.026498253	-0.003908904	0.03295559
4	0.057325684	0.008054568	-0.024977243	0.019409794

Table 10. *Optimal balance of 4 individual securities found from MPT*

	AMZN	PG	JPM	PLD
0	0.7290881744321444	0.2194274733708327	0.0317339537650592	0.0197503984319635
1	0.1166246964635128	0.2504026143417068	0.4250607619124834	0.2079119272822971
2	0.8981467701839371	0.0332089134841185	0.0175451379625771	0.0510991783693673
3	0.6073852442477203	0.0060461060834225	0.0489354062565848	0.3376332434122723
4	0.7008534358887363	0.0889681528692743	0.0036783632235842	0.2065000480184051

4. Conclusion

The optimization of risk-return ratio of a portfolio using Modern Portfolio was effective. It could create a portfolio that can perform better than the Dow Jones Industrial Average with 30 individual stocks. The CAGR of the optimal portfolio was 14.03% from 2017 to 2021 while the CAGR of Dow 30 was 11.70%. It also performed better than the 11 sectors from the GICS, having a CAGR of 26.65% for optimal portfolio while the S&P 500 had a CAGR of 15.02% from 2017 to 2021. The result was similar for the selected 4 sectors from the GICS, and this led to a conclusion that the best result was for a portfolio with around 10 individual securities, while it shows similar results for a portfolio with around 5 individual securities. The application of the forecasting LSTM model drastically improved the performance of Modern Portfolio Theory, as its CAGR recorded 88.03% while the equally balanced portfolio only had a CAGR of 23.62%.