# 國立臺灣科技大學資訊工程系

## 110 學年度第 1 學期專題研究

總報告

研究題目：Using LSTM to Predict Stocks Price and Apply Black-Litterman Model to Do Asset Allocation

組員


| | |
|---|---|
| B10604044 | 李明揚 |
| 學號 | 姓名 |
| 學號 | 姓名 |


指導教授：柯拉飛

中　華　民　國　111 年　1 月　20 日

# Using LSTM to Predict Stock Price and Apply Black-Litterman Model to Do Asset Allocation

**Group Member:** B10604044 李明揚, mail: s10309005@gmail.com
**Research lab name:** 網路協定與資源分配最佳化設計實驗室
**Professor:** Rafael Kaliski

## I. Introduction

In stock market, it is important to spread the risk. Therefore, I apply Black-Litterman Model in my work. Black-Litterman Model(BL Model) is one of the best way of Asset Allocation. By computing the rate of return of each stock as input, then we can get the configuration of asset allocation. According to the estimate of rate of return, this is the best combination. However, what is the method to get the estimate of rate of return? There are two methods. First, I used LSTM model to predict. Using the history stock price from 2018/3 to 2021/3 to train model, then predicting the whole month on April, 2021. Second, I used the target price from investment company/bank.

## II. Background Research

Before BL model created, The mainstream asset allocation model is the Markowitz Mean-Variance model. This model define how to diversify investment so as to minimize risks and maximize returns. This is the first time that mathematical statistics methods have been introduced into portfolio theory. It is based on the expected return and standard deviation of various portfolios. The following function can help investor get the optimal asset allocation:
max┬ω⁡〖ω^' μ-1/2 δω'Σω〗,
where μ is expected rate of return on asset, Σ is covariance matrix of asset, δ is investor risk aversion coefficient, and ω is weight of assets in the portfolio, and the best solution is:
ω=〖(δΣ)〗^(-1) μ.
Although Markowitz Model can get the best solution, it is still not suitable for real-world applications. Because this model is sensitive to expected rate of return on asset. For example, if the return rate of one asset A is increased by 1%, the Markowitz model originally calculated that it was a small-volume sale, but it would become a large-volume purchase. So this model is unstable, and the BL model later improved this problem.

## III. Problem Statement

In stock market, there are a lot of investment company such as some foreign bank, JPMorgan, UBS, Deutsche Bank, Nomura securities, etc., or native bank, Mega International Investment Trust Co.,Ltd., Yuanta Securities, etc. These investment companies will evaluate the company, may analyze the company's financial status and future prospects, and then write a report to give the company's stock a target price. Investors can evaluate whether to invest in this company based on these research reports. However, investment companies do not necessarily evaluate every company, so when investors want to know more about these companies, there will be less reference data. Therefore, deep learning can be used to help judge at this time. You can use LSTM to predict the stock price trend, and the LSTM trained with historical stock prices can effectively judge the future trend, and you can make up for the shortcomings of the target price of the investment company to invest.

Therefore, We have prepared 4 stocks, all of which have target prices obtained by the investment company's analysis of them. Then compare it with the future stock price predicted by LSTM. Then put the expected rate of return obtained by these two different methods into the BL model that deals with asset allocation, and compare two different portfolios to verify whether this investment method is feasible.

## IV. Proposed Solution

### A. *Black-Litterman Model*

Black-Litterman model is a mathematical model for portfolio allocation developed in 1990 at Goldman Sachs by Fischer Black and Robert Litterman, and published in 1992. It seeks to overcome problems that institutional investors have encountered in applying modern portfolio theory in practice. The model starts with an asset allocation based on the equilibrium assumption$_1$ and then calculate that allocation$_2$ by taking into account the opinion of the investor regarding future asset performance$_3$.
1.Assets will perform in the future as they have in the past 2.Using the formula below to calculate 3. Expected return on assets
The formula is:

$$E[R] = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}\,[(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q]$$

τ: scale factor, mean uncertainty, const
Σ: Covariance matrix of Excess rate of return$_1$, size N*N
P: Investor's view matrix, size K*N
Π: Implied equilibrium rate of return vector, size is N*1
Q: Point of view earnings vector$_2$, size K*1
Ω: Covariance matrix of opinion error, used to express the difference between the investor's opinion and the actual situation, size K*K
K: The number of view N: The number of asset

In my work, τ is 0.05, Σ is the covariance matrix calculated using the historical daily rate of return of the four stocks, the formula is as follows:

$$\Sigma = \delta * F + (1 - \delta) * S$$

,where S is identity matrix of the historical daily rate of return matrix, and F is a diagonal matrix which size is 4*4. Each row of matrix P represents a viewpoint, and each column represents a stock. In my work, I make the matrix equal in the two method. P is 4*4 identity matrix, this mean that computing the rate of return of the stocks without discussing the relationship between stocks. Π is none. Matrix Q is the rate of return for each stock based on two methods. In Report method, Q is [0.1969112, 0.10732538,

0.27472527, 0.10915493]T . In LSTM method, Q is [-0.0124535, 0.03225983, 0.25351432, 0.06221124]. The formula of matrix Ω is: τP^' ΣP.

### B. *Long Short Term Memory(LSTM)*

#### 1) *Data Collection*

I used yahoo finance API on python to collect 3 years stock data to train model, and use this model to predict April, 2021.

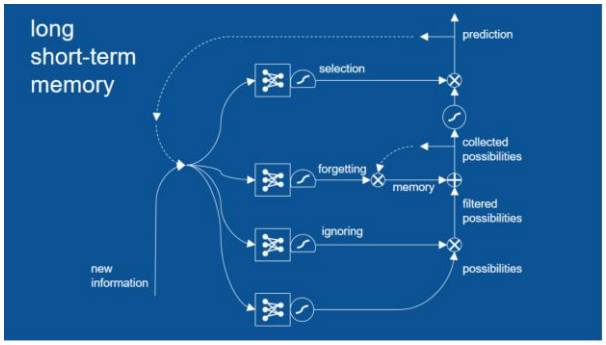
Finding the maximal stock price in the predicted data, then calculate the rate of return to apply it on BL model as a parameter matrix(Investor viewpoint matrix). Therefore, we can get the asset allocation from BL model.

#### 2) *Data Preprocessing*

I made data split into training data and testing data. First, the training data was normalized between 1 and 0; the purpose is that avoiding the data exceeding situation. Then, I made training data 20-day cycle, because 20-day is a month working days, it belongs mid-to-long term view. Training Label is daily open price.

#### 3) *What is LSTM?*

LSTM consists of four units, input gate, output gate, memory cell, forget gate. Input gate can decide whether to refer to during data input. Output gate can control whether to output the value, if not, output 0. Memory cell can memorize the value for the next cell. Forget cell can decide whether to forget the value.




Brandon Rohrer(2017)

#### 4) *Building Model*

I used keras “Sequential” to build a model. First, I added a “LSTM” layer, units = 100, then I added a Dense layer to make output dimension=1. In the end, compiling the model, using ‘Adam’ optimizer, loss function is ‘mean_square_error’. We use it to fit data, epochs = 30, batch size = 16.

#### 5) *Why is LSTM?*

I compare this model with SVR(Support Vector Regression). SVR uses SVM to do regression analysis. SVM is good to solve classification problem, but in regression, this algorithm is very sensitive to parameters and kernel function. Because the parameter ‘c’ in kernel function ‘rbf’ which can decide the difficulty of decision function. If c is small, the decision function would be simpler; however, the accuracy will be unstable. Moreover, the parameter ‘gamma’ defines the influence of a single training reaches, low values means ‘far’ high values means ‘close’. Therefore, SVR is very sensitive to gamma, if gamma is too large, the model would overfitting. But if gamma is too small, the model cannot capture the complexity of the data

So the performance of SVR is worse than that of LSTM. We look at it from the perspective of root mean square error.

We used the data from March, 2018 to March, 2021 to train model. Root mean square error (RMS)formula is: sqrt(mean((real_data - predicted_data)^2))

RMS of LSTM: 1.72, RMS of SVR: 10.14. Therefore, we can find that LSTM model is more better than SVR in this view.

Figure 1 and Figure 2 are stock 2603, Figure 1 is using LSTM to predict, and Figure 2 is using SVR to predict, we can compare the forecast for April.

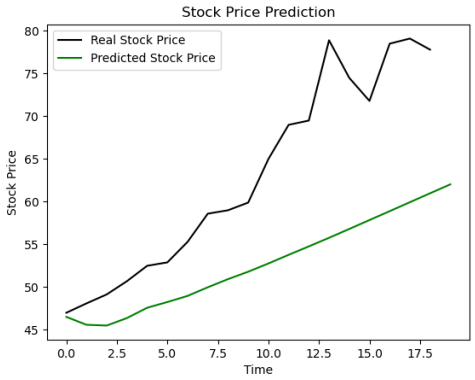



Figure 1

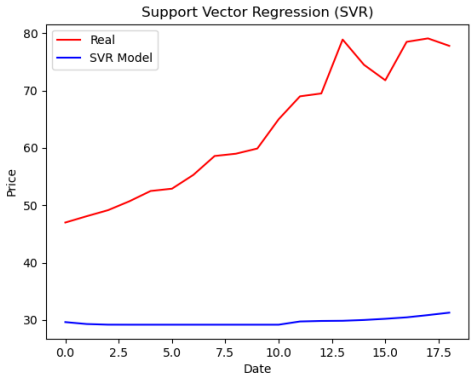



Figure 2

### C. *Report of Investment Company*

I got the target price of the stock from 4 credible investment companies. I used it to calculate the rate of return and apply it on BL model as a parameter matrix(Investor viewpoint matrix). Then, we could get a asset allocation that is different from the using LSTM. Then, we discuss the rate of return in April, 2021.

| Stock | 2002 | 2330 | 2603 | 2881 |
|---|---|---|---|---|
| Investment Company/Bank | 兆豐 | 瑞士信貸 | 兆豐 | 宏遠 |
| Target Price | 31.0 | 650.0 | 58.0 | 63.0 |

Source: 籌碼 K App

## V. Simulation

### A. *BL Model Asset Allocation*

I used python API pypfopt.black_litterman to build BL Model. The calculation method I have explained in the second point. We only need to compute Q matrix. The Q matrix uses the obtained rate of return as input. The following are the configurations obtained after applying the BL model respectively:

| Stock ID | 2002 | 2330 | 2603 | 2881 |
|---|---|---|---|---|
| 4/01 | 26 | 598 | 47 | 57.6 |
| 4/29 | 39.5 | 609 | 77.8 | 64.9 |
| Rate of Return(%) | 51.9 | 1.5 | 65.5 | 12.7 |

LSTM

| Stock ID | 2002 | 2330 | 2603 | 2881 |
|---|---|---|---|---|
| Config | -0.08 | -0.49 | 1.5 | 0.07 |

Report

| Stock ID | 2002 | 2330 | 2603 | 2881 |
|---|---|---|---|---|
| Config | 0.74 | 0.015 | 0.195 | 0.05 |

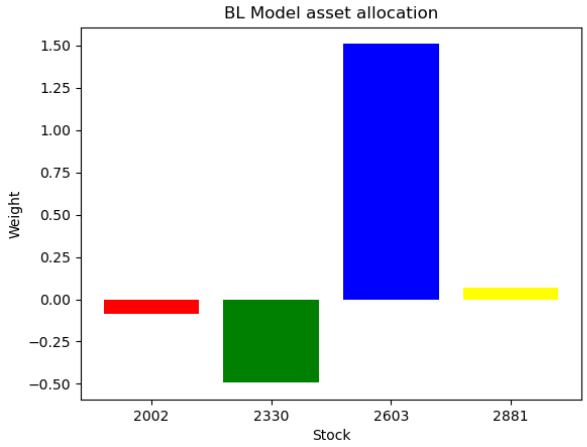



Figure 3

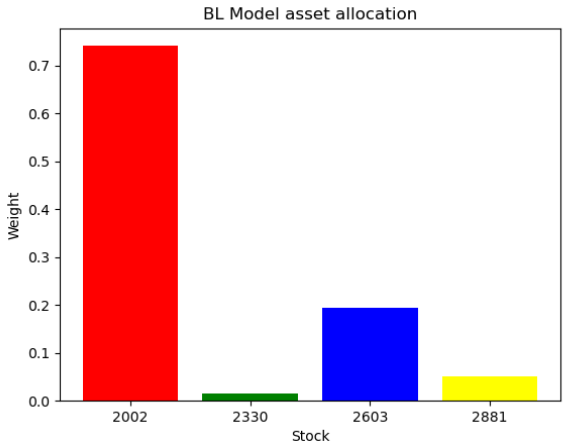
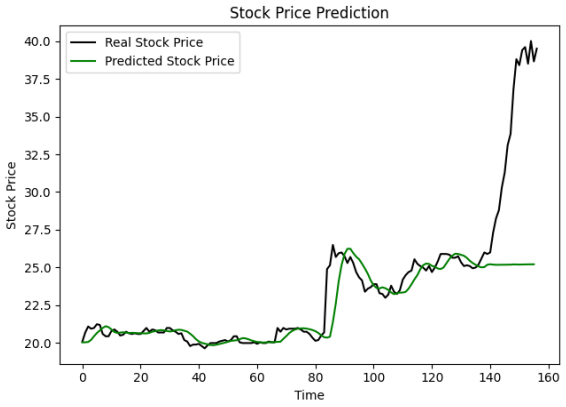


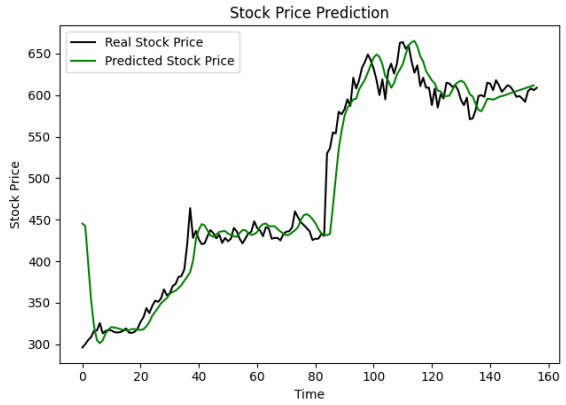

Figure 4

*B. Result of LSTM*

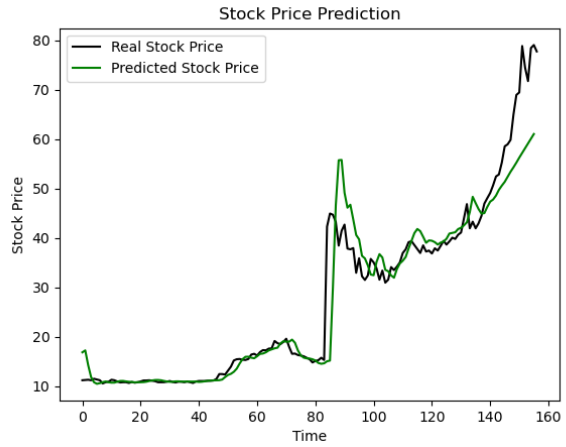
Following figures are the prediction of four stocks.



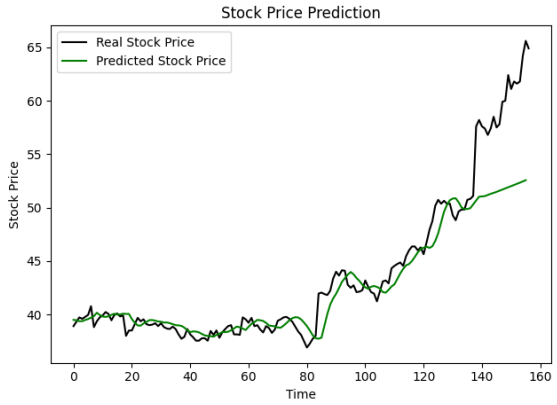

This is the prediction of stock 2002 中鋼, the predicting rate of return is -1.5%.




This is the prediction of stock 2303 台積電, the prediction rate of return is 3.2%.




This is the prediction of stock 2603 長榮, the prediction rate of return is 25.4%.




This is the prediction of 2881 富邦金, the predicting rate of return is 6.2%.

### C. *The Rate of Return*

We calculate the stock price on the last day of April to calculate the rate of return, We buy all stocks when stock market opened on April 1st, and then sell stocks on April 29th . The formula of Rate of Return is:

$$\text{Rate of Return} = \frac{Sell\ Price - Purchase\ Price}{Purchase\ Price} \times 100\%$$

The stock price(open price) of April:

Assume we have NTD20,000,000 initial funds. And we can long(Buy before Sell) or short(Sell before Buy) stock. Let's calculate the rate of return of each method.

*1) LSTM*

| Stock ID | 2002 | 2330 | 2603 | 2881 |
|---|---|---|---|---|
| | Short | Short | Long | Long |
| Money | 800,000 | 4,900,000 | 15,000,000 | 700,000 |
| Earn | -415,200 | -73,500 | 9,825,000 | 88,900 |

Total profit: (-415,200) + (-73,500) + 9,825,000 + 88,900 = 9,425,200

The Rate of Return: Total profit / Initial Funds = 9,425,200 / 20,000,000 = 47.13%

*2) Report*

| Stock ID | 2002 | 2330 | 2603 | 2881 |
|---|---|---|---|---|
| | Long | Long | Long | Long |
| Money | 14,800,000 | 300,000 | 3,900,000 | 1,000,000 |
| Earn | 7,681,200 | 4,500 | 2,554,500 | 127,000 |

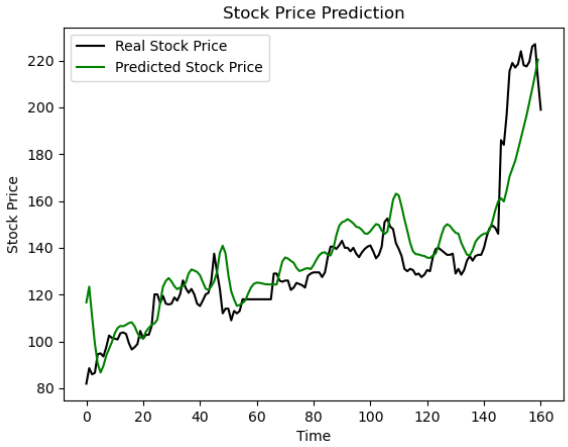
Total profit: 7,681,200 + 4,500 + 2,554,500 + 127,000 = 10,367,200

The Rate of Return: Total profit / Initial Funds = 51.84%

## VI. Conclusion

We can find that the rates of return of the two are very close. The research report of the investment company is the conclusion obtained after a series of precise analysis of company. The target price is set by a very professional person who specializes in accounting, international situations, economic conditions, business cycles, etc. To give an evaluation. However, they don't analyze each company to customers. For example, when a new industry is just starting, investment companies have no data to analyze, such as low-orbit satellites, stock 5222 全訊 is a high-tech company that develops and manufactures RF and microwave communication chips and modules. Their products are necessary materials for future communication satellites. In the future, the main 6G protocol will launch a large number of low-orbit satellites, so the company's performance is bound to be greatly improved. However, no investment company has written an investment report for such a promising company. Only one report was released in early November, 2021, but the company's stock price had risen to the end at that time.

But if we use LSTM to predict stock price for reference. I use November, 2018 to November 8, 2021 history stock price as dataset, using November, 2018 to March 1, 2021 as training data to train model, then I use November, 2018 to November 8, 2021 as testing data to predict the following stock price on November, 2021.Figure 3 is the prediction of stock 5022 using LSTM. The reason why I decide to predict stock price after November 9, 2021 is it is because a investment company release its investment report on November 9, 2021, the target price is NTD210. Therefore, we can compare the once again, the maximum price using LSTM to predict is 220.4, this price is more close than report of investment company (The real max price is 234). Therefore, LSTM is a nice tool to help us to invest, it is not the only reference we can refer, but it can tell us a direction, whether this stock will rise or not.




Both methods can tell us the stocks will rise or not, but how to decide the number of stocks we buy to be a portfolios. Black-Litterman model is a helpful tool. We can see using the result of LSTM prediction and report of investment company to execute BL model, their performence of portfolios are better than Taiwan Stocks Weighted Index, the rate of return of LSTM is 47.13%, and that of report is 51.84%; but the rate of return of weighted index is 6.22%. Generally speaking, if our portfolio beats the rate of return of the weighted index, we will say it is a successful investment.

## VII. Reference

[1] Y. -J. Yang, L. -L. Zhang, Y. Niu and O. -J. Zhang, "A Brief Review of Two Classical Models for Asset Allocating," 2020 16th International Conference on Computational Intelligence and Security (CIS), 2020, pp. 159-163, doi: 10.1109/CIS52066.2020.00042.

[2] M. Gan, L. Chen, C. -Y. Zhang and H. Peng, "A Self-Organizing State Space Type Microstructure Model for

Financial Asset Allocation," in IEEE Access, vol. 4, pp. 8035-8043, 2016, doi: 10.1109/ACCESS.2016.2626720.
[3] M. Day, J. Lin and Y. Chen, "Artificial Intelligence for Conversational Robo-Advisor," 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2018, pp. 1057-1064, doi: 10.1109/ASONAM.2018.8508269.
[4] R. Zhang, C. Huang, W. Zhang, S. Chen, “Multi Factor Stock Selection Model Based on LSTM”, 2018 International Journal of Economics and Finance, Vol. 10, No. 8(2018), doi: 10.5539/ijef.v10n8p36
[5] Xiaoping Wang and Baozhong Qu, "The study of Black-Litterman-Inflation asset allocation model," 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2011, pp. 1760-1762, doi: 10.1109/AIMSEC.2011.6010725.
[6] How Recurrent Neural Networks and Long Short-Term Memory Work(2017, June 27) Retrieved from: https://e2eml.school/how_rnns_lstm_work.html
[7] Portfolio Optimization Technology Black-Litterman Model(2020, November 13) Retrieved from: https://www.yuanta-etfadvisor.com/article/3f1b3088-0e59-475c-a6e9-6610282648be
[8] Modern Portfolios Theory(2021, September 10) Retrieved from: https://www.investopedia.com/terms/m/modernportfoliotheory.asp
[9] Machine Learning: Kernel function(2018, Jun 21), Retrieved from: https://chih-sheng-huang821.medium.com/%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-kernel-%E5%87%BD%E6%95%B8-47c94095171
[10] Five Methods to download stock price history(July 14), Retrieved from: https://medium.com/%E5%B7%A5%E7%A8%8B%E9%9A%A8%E5%AF%AB%E7%AD%86%E8%A8%98/5%E7%A8%AE%E6%8A%93%E5%8F%96%E5%8F%B0%E8%82%A1%E6%AD%B7%E5%8F%B2%E8%82%A1%E5%83%B9%E7%9A%84%E6%96%B9%E6%B3%95-766bf2ed9d6

## VIII. Appendix

Source code:
https://reurl.cc/AK7xQQ