

Performance Analysis for Portfolio Construction and Stock Price Prediction of Social Media Platforms

Xintong Lin^{1, *, †}, Xuepu Tan^{2, *, †}, Xuening Yang^{3, *, †}

¹Department of Computing, Mathematics, University of Hong Kong, Hong Kong, 999077, China

²The College of Liberal Arts and Sciences, Arizona State University, Tempe, AZ 85281, USA

³Beijing-Dublin International College, Beijing University of Technology, Beijing, 100124, China

*Corresponding author: leaf99lin@gmail.com, xtan40@asu.edu, 13220137372@sina.cn

†These authors contributed equally.

Abstract. The stock price prediction and portfolio construction are very important in the application of quantified investments. Risk control and stock price prediction are two challenge problems that we want to solve. In this paper, two main steps are processed to quantify the risk and predict the stock price. Especially, the first step is portfolio optimization, which is based on the Mean-Variance method. The second step is the stock price prediction, which is designed based on the Arima-based method. The results show that the stock “TWTR” and the stock “WB” have the above 90% ratio of the stock combination, which shows that the “Twitter” and “Weibo” social media corporations have a higher expected return with a lower risk. In addition, the best ARIMA model used to fit our data will be the one with parameters $p=4$, $d=0$, $q=3$. Our research has great significance in the application of quantified investments.

Keywords: component; Portfolio Construction; Stock Price Prediction; Social Media Platforms; ARIMA Model

1. Introduction

Portfolio optimization aims to explore the best combination of investments to obtain the biggest return. The main portfolio characteristics used in portfolio optimization suggested by Markowitz [1] are the expected return and the risk. In contrast to the expected return, the specification of the portfolio risk appears to be a more complicated task [2].

In Markowitz’s portfolio theory [1], the variance is taken as a risk measure. Then optimal portfolios are constructed by minimizing the variance for a given level of the expected return or maximizing the expected return for a given value of the variance. However, these optimization problems take into account only one characteristic of the portfolio while the second is fixed. Especially, the variation of the stock price in the social media platforms market seems stable. Therefore, the investment strategy for the social media platforms market may be specific. It is very important for us to find a suitable measurement standard to minimize the risk of this investment is very important.

Many researchers proposed some methods to measure the risk. For example, Sharpe [3, 4] defined the Sharpe ratio (SR) as a ratio of the expected portfolio return to the standard deviation. Please note that the optimal portfolio in the sense of maximizing the SR belongs to the efficient frontier in the case without a risk-free asset, which can be obtained as a solution to Markowitz’s optimization problem. Krokhamal et al. [5] proposed the downside risk measures that depend only on the positive values of the return’s loss function or negative values. The risk theory has been pushed that the quantile-based measures are well-suited functions to quantify risk [6-13].

The pioneering work of Markowitz on the mean-variance (MV) portfolio optimization procedure is the milestone of modern finance theory for optimal portfolio construction, asset allocation, and investment diversification [14].

We hope to optimize the investment portfolio by constructing the mean-variance so that investors can diversify some risks. The optimized portfolio approach provides investors with a way to reduce asset-related risks without necessarily reducing their expected rate of return. Compared with a simple

equal-weighted portfolio, this improves the trade-off between risk and return because a lower level of risk is achieved at the same level of return. In addition, we have also realized the investment portfolio return forecast for the next 21 days. This helps us to observe the return range of the optimized investment portfolio more intuitively.

The main reason why the latter has traditionally been preferred is that the semicovariance matrix, contrary to the covariance matrix, is endogenous (its terms change every time the portfolio weights change); therefore, optimization problems that use it are intractable. A solution to this problem is provided in Estrada [15], who proposes a heuristic that computes the elements of the semicovariance matrix with respect to when the single assets, and not the portfolio as a whole, underperform the benchmark. This procedure yields an exogenous matrix that well approximates the semicovariance matrix and can be used for portfolio optimization. It has been pointed out in Estrada [15] that it would be deceiving to compare the results obtained from optimization procedures that employ the variance with those obtained from procedures that employ the semivariance using an index based on either mean-variance or mean-semivariance efficiency. This is because, by construction, mean-variance optimization will appear to perform best when using a mean-variance performance measure (such as the Sharpe ratio). In contrast, mean-semivariance optimization will appear to be the best when using a mean-semivariance performance measure (such as the Sortino ratio).

The contributions of this paper include 1) The MV-based method is used to maximize the sharp ratio portfolio; 2) The ARIMA Model is used to predict the price trend of the selected stocks. 3) The comparison of ARIMA and machine learning-based methods is given to show the method's performance for stock price trend prediction.

The rest paper includes Section 2 introduces the methods; Section 3 introduces the results and discussion; Section 4 introduces the conclusion.

2. Method

In this paper, two steps are designed to achieve the goal that we can obtain the best return from the stock market. The two steps include “Portfolio Optimization” and “Stock Price Prediction”. Especially, portfolio optimization is used to get the best strategy, and the stock price prediction is used to validate the effect of the above strategy. For example, if the strategy is useful, the return would be good, and the stock price prediction may get a good result with high accuracy.

2.1 Data Preparation

Five stocks, including “Twitter (TWTR)”, “Facebook (FB)”, “Snap (SNAP)”, “Zoom Video (ZM)” and “Weibo Corporation (WB)” are selected from the Yahoo Finance website. The five stocks can represent the rule of the social media market because the above five stocks are the typical stocks in this market. The variation of these stocks following the time series can be found in figure 1.

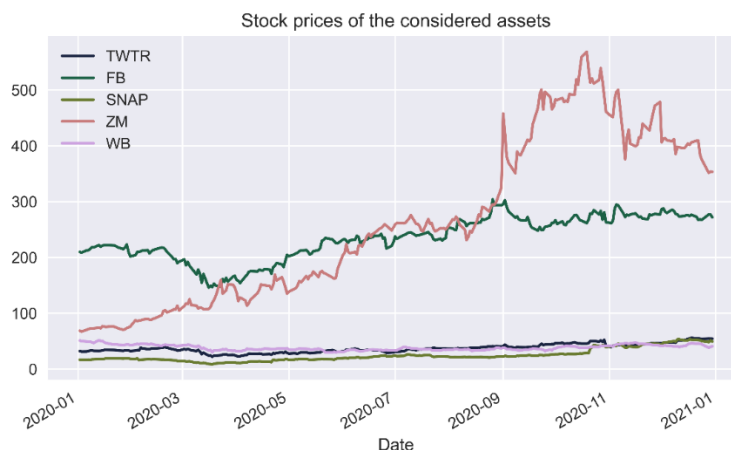


Figure 1. The stock price variation following time series.

The y-axis means the stock price; the x-axis means the time series of the stock market. Five stocks are shown in different color lines in legend. Obviously, from figure 1, we can find that all the variation of stock prices is small except for the “ZM”. In summary, the data selected has good stability. This step can decrease the risk of the market, which could decrease the difficulty of the problem.

2.2 Portfolio Optimization

In this paper, the best return should be under the smallest risk. Therefore, risk quantification is very important. For example, the benchmark below which volatility is considered to be downside volatility depends on the investor's preferences, and it does not necessarily coincide with the mean portfolio return. If returns are perfectly symmetrically distributed, the investor requires to directly target the volatility below the benchmark. Finally, the measurement by the downside deviation can be calculated by the formula (1).

$$\sigma_{\beta} = \sqrt{\frac{1}{T} \sum_{t=1}^T [\text{Min}(r_t - B, 0)]^2} \quad (1)$$

T means the whole-time series, and r represents the returns; the investor proposes a benchmark B set.

2.3 Stock Price Prediction

In this paper, the Arima method is used to predict the stock price. In forecasting form, the best model selected can be expressed as formula (2),

$$Y_t = \phi_t Y_{t-1} + \theta_0 + \epsilon_t \quad (2)$$

where $\epsilon_t = Y_t - \hat{Y}_t$ is the difference between the actual value and the forecast value of the series.

We iterate through a non-trivial number of combinations of (p, d, q) orders to find the best ARIMA model to fit our portfolio returns. We use the AIC to evaluate each model. The lowest AIC wins.

3. Results and Discussion

In this paper, the results include two parts: portfolio optimization results and stock price prediction results.

3.1 Results of Portfolio Optimization

The daily return of considered assets is counted statistically shown in figure 2.

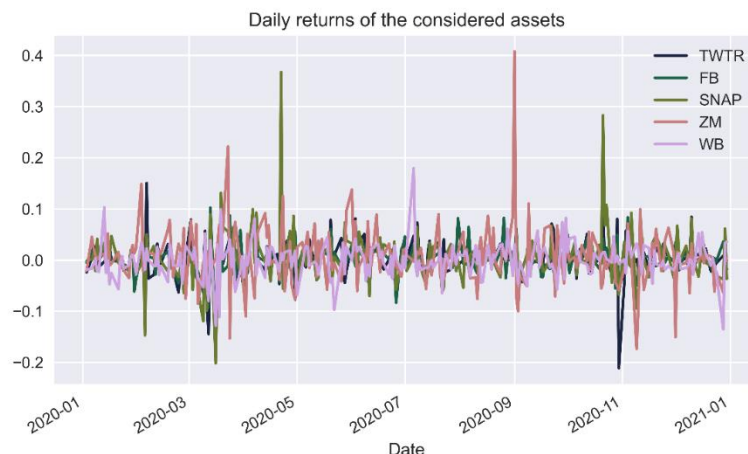


Figure 2. The Distribution of daily returns.

From figure 2, the results show that the daily returns obey the normal distribution. And the results of the expected return have two parts: the maximum sharp ratio and the minimum volatility. The results can be found in table 1.

Table 1. Results of Selected stocks RATIO (Portfolio)

Stock Name	Maximum Sharp ratio	Minimum Volatility
FB	5.99%	0.98%
SNAP	2.38%	51.16%
TWTR	35.34%	0.22%
WB	56.09%	18.09%
ZM	0.19%	29.55%

In table 1, the first column shows the membership of stock combination, which includes “FB, SNAP, TWTR, WB, and ZM”. The second column shows the results of the maximum sharp ratio. The third column shows the results of minimum volatility. Especially, the maximum sharp ratio means the maximum expected return per risk unit. Besides, the minimum volatility means the minimum risk we want to accept.

The maximum sharp ratio can represent the best performance of the model we selected. “TWTR” and “WB” have the above 90% ratio of the stock combination, which shows that the “Twitter” and “weibo” social media corporations have a higher expected return with a lower risk. The portfolio optimization process can be found in figure 3.

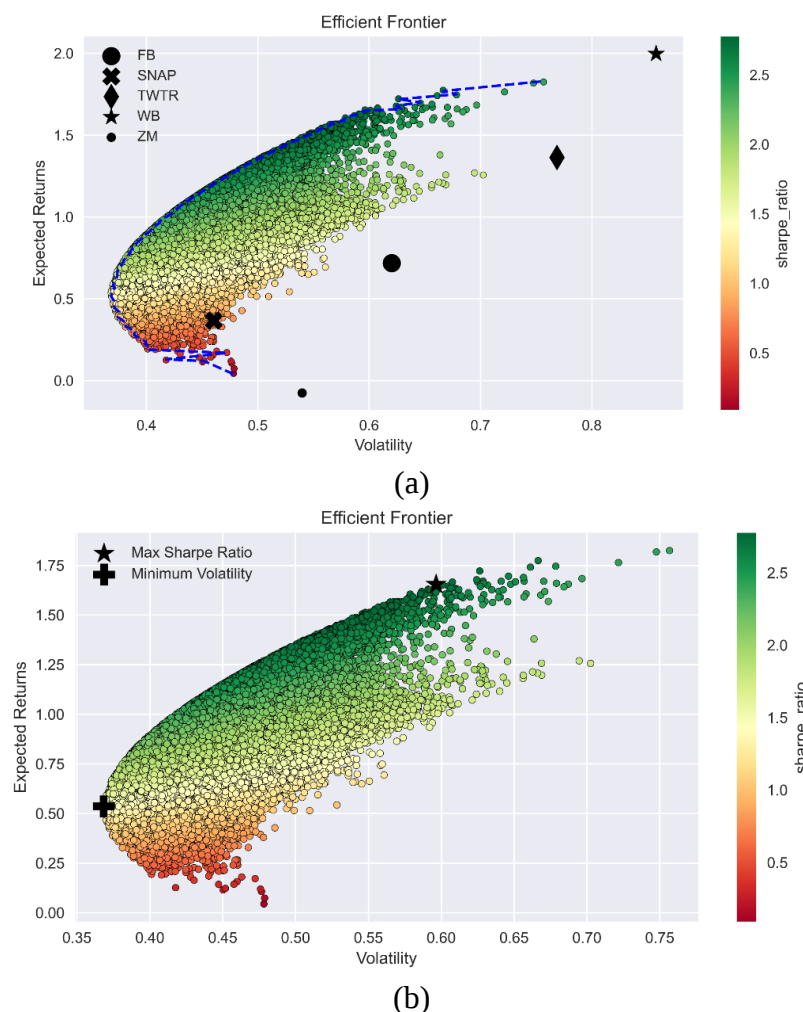


Figure 3. The process of portfolio optimization.

In figure 3, subfigure (a) shows the scatter distribution of “Efficient Frontier”, and subfigure (b) shows the comparison of maximum sharp ratio and minimum volatility. Especially, subfigure (a) shows that the “TWTR” and “WB” have the highest expected return ratio. However, they also have the highest risk ratio. After considering the balance of return and risk, we use the maximum sharp ratio as our optimization strategy.

3.2 Results of Stock Price Prediction

The results of stock price prediction are shown in Figure 4 and Table 2.

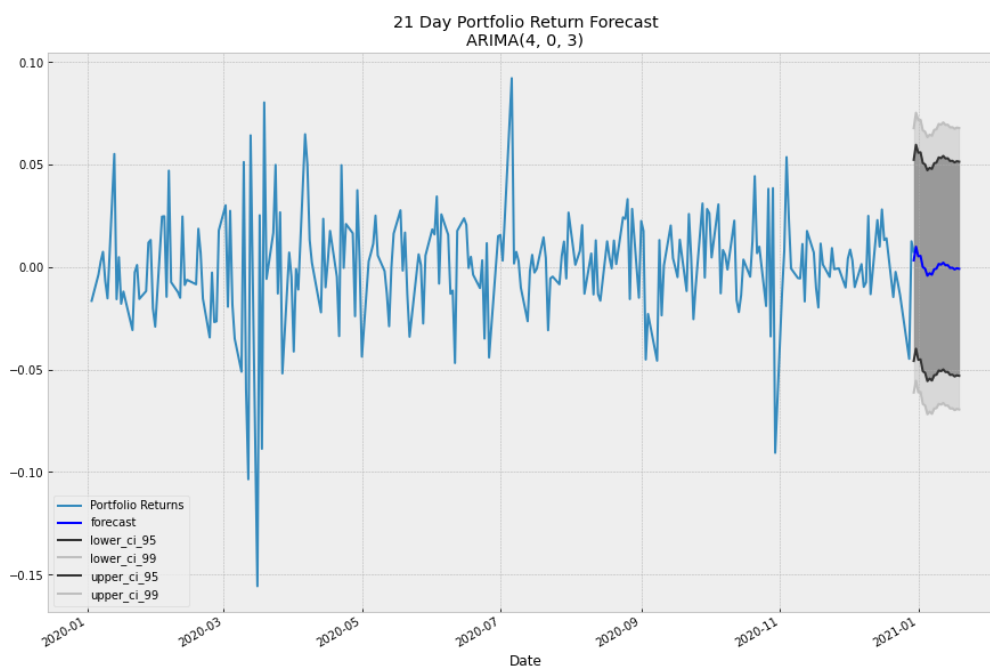


Figure 4. 21 Day Portfolio Return Forecast

The result of the 21-day portfolio return forecast is counted statistically shown in figure 4 in the grey part, including the forecast value and the 95 percent and 99 percent confidence interval.

From figure 4, the results show that daily returns of the portfolio we constructed in the following 21 days are waving around the axis, with both sides of the 95 percent confidence interval has an absolute value of less than 6 percent. Although there is a fluctuation there, the movement is statistically slight. The details of the first 5 days prediction are shown in table 4. The returns are positive for the first day and negative on day 5.

Table 2. The first 5 days forecasting

Day	forecast	Lower_ci _95	Lower_ci _99	Upper_ci _95	Upper_ci _99
2020-12-30	0.0033	-0.0458	-0.0612	0.0524	0.0678
2020-12-31	0.0100	-0.0398	-0.0554	0.0597	0.0754
2021-01-01	0.0053	-0.4525	-0.0611	0.0558	0.0717
2021-01-02	0.0055	-0.0450	-0.0609	0.0560	0.0719
2021-01-03	-1.2e-5	-0.0508	-0.0668	0.0508	0.0668

And the ARIMA model fits best with our original data is illustrated in Table 3. The result shows that the best ARIMA model to fit our data will be the one with parameters $p=4$, $d=0$, $q=3$. The model wins because of its AIC equaling 14182.319, which is the lowest among all the iterations. It should

be no surprise that the best model has a differencing of 0 since the first differences of log prices are used to calculate the stock returns.

4. Conclusion

This paper first used the mean-variance model to construct a portfolio with the maximum sharp ratio, which perfectly balanced the expected return with the portfolio's risk. According to the model we use, the expected return of the portfolio should be around 1.65%. Then, we predict the portfolio's return of the next following days with the ARIMA model. The prediction shows that returns of the portfolio are below 1%, which is also relatively low. The development of social media is fast and expected to be optimistic. However, investing in such a portfolio with the maximum sharp ratio does not seem to be a good choice. For people who are optimistic about developing social media platforms, if they are willing to take a certain risk, then investing in a single stock is expected to be a better choice for them. If they are typical risk-averse investors, then they should invest in market portfolios that provide higher yields, such as S&P 500.

References

- [1] Markowitz H. Portfolio selection[J]. 1959.
- [2] Bodnar T, Zabolotskyy T. How risky is the optimal portfolio which maximizes the Sharpe ratio? [J]. AStA Advances in Statistical Analysis, 2017, 101(1): 1-28.
- [3] Sharpe W F. Mutual fund performance[J]. The Journal of business, 1966, 39(1): 119-138.
- [4] Sharpe W F. The sharpe ratio[J]. Journal of portfolio management, 1994, 21(1): 49-58.
- [5] Krokmal P, Zabarankin M, Uryasev S. Modeling and optimization of risk[J]. Handbook of the fundamentals of financial decision making: Part II, 2013: 555-600.
- [6] Acerbi C, Tasche D. Expected shortfall: a natural coherent alternative to value at risk[J]. Economic notes, 2002, 31(2): 379-388.
- [7] Artzner P, Delbaen F, Eber J M, et al. Coherent measures of risk[J]. Mathematical finance, 1999, 9(3): 203-228.
- [8] Acerbi C, Tasche D. Expected shortfall: a natural coherent alternative to value at risk[J]. Economic notes, 2002, 31(2): 379-388.
- [9] Alexander G J, Baptista A M. Economic implication of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis[J]. Journal of Economic Dynamics and Control, 2002, 26(7-8): 1159-1193.
- [10] Alexander G J, Baptista A M. A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model[J]. Management science, 2004, 50(9): 1261-1273.
- [11] Kilianová S, Pflug G C. Optimal pension fund management under multi-period risk minimization[J]. Annals of Operations Research, 2009, 166(1): 261-270.
- [12] Bodnar T, Schmid W, Zabolotskyy T. Minimum VaR and Minimum CVaR optimal portfolios: estimators, confidence regions, and tests[J]. Statistics & risk modeling, 2012, 29(4): 281-314.
- [13] Bodnar T, Schmid W, Zabolotskyy T. Asymptotic behavior of the estimated weights and of the estimated performance measures of the minimum VaR and the minimum CVaR optimal portfolios for dependent data[J]. Metrika, 2013, 76(8): 1105-1134.
- [14] Bai Z, Liu H, Wong W K. Enhancement of the applicability of Markowitz's portfolio optimization by utilizing random matrix theory[J]. Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 2009, 19(4): 639-667.
- [15] Estrada J. Mean-semivariance optimization: A heuristic approach[J]. Journal of Applied Finance (Formerly Financial Practice and Education), 2008, 18(1).