Stock Analysis and Portfolio Optimization Using modified ARIMA model

Dr. M N Nachappa[1], SagarJain[2],Shaphrang Lyndoh[2],Rishab Gandhi[2]

[1]Professor &Head,Department of CS& IT, Jain(Deemed-to-be-University

[2]Scholar, Department of BCA, Jain (Deemed-to-be-University)

Abstract–Stock analysis and Portfolio Optimization are some of the common terms in finance today. Historically, stock markets have proven to be the best asset class for generating high returns &achieving your financial goals. But with high returns comes high risk and if not managed properly can wipe out your entire investment. According to research more than 90% of investors lose money in the stock market and majority of them are retail investors, who generally invest based on their intuition. So, keeping this in mind this project is developed to help retail investors in their investment journey.In this project we aim to use the capabilities and power of python along with streamlit to analyze the stock data and try to make a general prediction of the stock price in the near future and also create & optimize our portfolio to maximize our returns for a given amount of risk**.** One must understand that stock picking is an art and not science and there is no single strategy that works all the time, therefore, the topic of portfolio optimization and stock price prediction is extremely broad, and there are many techniques and approaches that have evolved and will keep evolving.

Introduction

For many years, stock prediction always has drawn attention to the development of intelligent trading systems. There are substantial benefits that can be gained from stock prediction for security selection and quantitative investment analysis. And with the enormous growth of financial data in volume and complexity, increased processing capability of computers and evolution of machine-learning algorithms allow us to extract patterns from data processed.

There are many models and techniques used for prediction but in this project we use ARIMA model to make a general forecast of the likely trend the stock price is going to follow in the near future. ARIMA models are known to be efficient in time series forecasting. One important thing to note here is that, no matter what, it's impossible to predict and time the stock markets every time. But short -term prediction give us a general behavior of the stock.

As stated already that historically, stock markets in the long run have outperformed other investment options and proven to be the best asset class for generating high returns on investments. But these returns are not fixed and are subject to various 12

Objectives - Stock analysis and portfolio optimization is a project implemented using the python programming language with the intent of aiding investors in the analysis of securities and helping them in optimizing their portfolios.

Literature Review

The most remarkable work in the field of portfolio optimization was done by Nobel Laureate Harry Markowitz. Although, many other models have also evolved since then. According to the Markowitz, Risk and Reward are two aspects of investment considered by investors. The expected return may vary depending on the assumptions. If the investor wants high return, he has to take higher risk. Modern portfolio theory is the contribution of Harry Markowitz, which assists the investors how to diversify the risk. Essentially, MPT is an investment framework for the selection and construction of investment portfolios based on the maximization of expected returns of the portfolio and the simultaneous minimization of investment risk.

William Sharpe (1963) studied Markowitz's research and worked on simplifying the calculations in order to develop a practical use of the model.Under SIM(Single index model), ranking of the stocks (from highest to lowest) is done on the basis of their excess return to beta ratio to construct an optimal portfolio.

Modern portfolio theory is based on historical data and the model is extremely sensitive to inputs. The inputs are - expected return and expected volatility which are calculated as mean of past data. Many implementations of Markowitz Portfolio Optimization use the historical meanof daily return(usually one year period or more) and the variance of the past data.

However, the mean return of previous year may not be a reliable measure for the return of the stock. It may be that due to some factors the stock might have exhibited an abnormal behavior. A good company may have had a bad year or a bad company has shown a price movement that is suspicious.

CAGR (Compounded Annual Growth Rate)

To overcome these short-term effects, we use CAGR as a measure of expected return of the stock. This is what analysts usually use to estimate the long term performance of a stock and good stocks with good CAGR are generally expected to perform better over the long run. Ex-HUL has given a CAGR of 20% over the last 20 years, therefore, we can use this as a reference that if the same trend continues in the future we can expect HUL to give 20% compounded return every year.

VaR (Value at Risk)

The risk in Markowitz model is variance which is used as a measure of risk. Variance has a major drawback i.e. it considers the movement of stock in both the directions (positive and negative) but positive side is favorable to us and the negative side is the one which investorsdon't like. Therefore, we use VaR which is a quantitative and synthetic measure of risk.

A portfolio having 2% one-week VaR of 1%means that there is a 2% chance that the asset will decline by 1% within a single week. This is a better measure of risk than variance.

Methodology

The entire application has been implemented in python usingStreamlit. Streamlit is an open-source Python library that allow us to create and share data science and machine learning applications. The streamlit application takes stock ticker as input and returns the results in the form of interactive easy to understand visual charts.

- **ARIMA Model –** ARIMA stands for Auto Regressive Integrated Moving Average. There are two types of ARIMA models – Seasonal ARIMA and Non – Seasonal ARIMA. Since, stock data is usually Non – Seasonal we will be using the Non – Seasonal ARIMA. It takes 3 parameters (p,d,q) where p,d& q are non – negative positive integers.

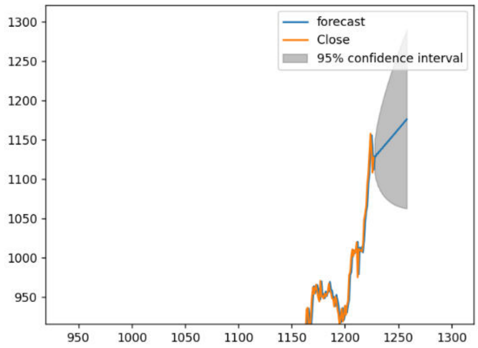
Parts of ARIMA Model

- **AR(p) – Autoregression** – A regression model that utilizes the dependent relationship between a current observation and observations over a previous period.
- **I (d) –Integrated** – Differencing of observations (Subtracting an observation from an observation at the previous time step) in order to makethetime series stationary.
- **MA (q) -Moving Average** – A model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations.

When fitting an ARIMA model, we aim to find the values of our parameters p,d and q which optimise or minimise a certain metric of interest. There are many methods to achieve this goal and yet the correct parametrization of ARIMA models can be a tedious process that requires statistical expertise and time. In this tutorial, we hope to overcome this issue by writing a grid search algorithm in python to select the optimal parameter values for our ARIMA(p,d,q) time series model.

The use of a "grid search" is to iteratively explore different combinations of parameters. For each combination of parameters, we fit an ARIMA model with the SARIMAX() function and assess its overall performance. Once we have explored the entire domain of parameters, our optimal set of parameters will be the one that yields the best performance for our criteria of interest. In this scenario, our criteria of interest is Akaike information criterion (AIC). The AIC measures how well a model fits the data while taking into account the overall complexity of the model. We are therefore interested in finding a model that returns the lowest AIC value.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 05 | May - 2021 ISSN: 2582-3930



- **Portfolio Optimization–**List of stocks and initial investment amountis taken as input from the user to create a portfolio. The user must also input his allocation for each stock which will be compared with the optimized allocation.

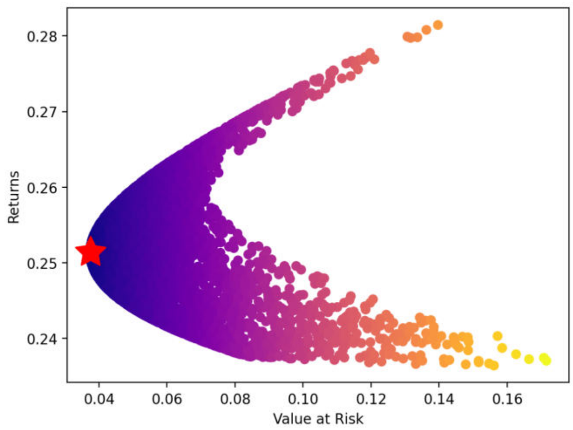
We run ten thousand simulations of random portfolios,byassigning random weights to securities in each portfolio and then calculateits CAGR and Value at Risk. Finally, the portfolio with minimum Value at risk that gives thebest CAGR is selected. The users can also enter their expected return if they want more returns and the program will give the most optimal portfolio with minimum VaRfor that return.

$$CAGR = [(Ending\ value\ /\ Beginning\ value)^{1/No.\ of\ years} - 1] * 100\%$$

For the calculation of VaR we scale the portfolio standard deviation by the square root of the "days" value, then subtract the scaled standard deviation, multiplied by the relevant "Z value" according to the chosen value of "alpha" from the portfolio daily mean returns which have been scaled linearly according to the "days" value. We have chosen 252 trading days (the number of trading days in a year) and an alpha of 0.05 with a 95% confidence level.

Results

We then plot the allocations on a chart showing CAGR vs VaR. From this plot we can find the optimal portfolio and extract its allocation to create our portfolio.

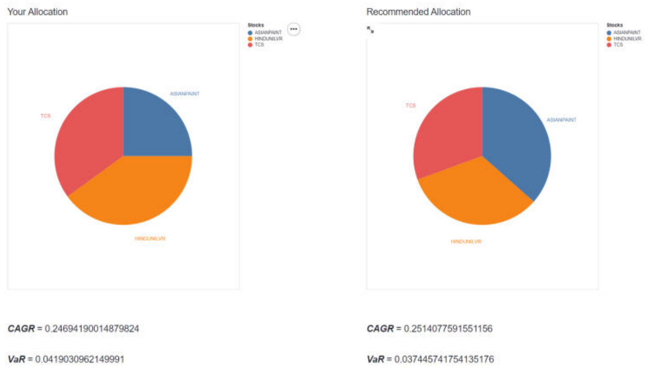


The minimum VaR Optimal portfolio is compared to the user input stock allocation.

The input allocations for ASIANPAINT, HINDUNILVR & TCS were 25%, 40% and 35% and the recommended allocation after portfolio optimization were 36.51%, 32.92% & 30.57% respectively.

It can be clearly seen that the optimized portfolio gives better CAGR (returns) and also has less VaR (risk) than the user input portfolio allocation. Though the CAGR doesn't seem very high but in the long run even 1% compounded annually could add a lot of value to ourinvestments.

Volume: 05 Issue: 05 | May - 2021 ISSN: 2582-3930



Conclusion– This project shows how streamlit can be used to create a stock analysis tool to aid investors in their investment process. The ARIMA model can be used to gain insights into to the future behaviour of the stock prices and portfolio optimization can be used to construct an efficient portfolio using the CAGR and VaR approach based on your investment choices. However, in markets there are no strategies that work all the time and hence, other risk factors should also be considered. The limitations of this project is that the portfolio stocks should have a long history of returns so that the expected CAGR could be calculated and used effectively. In future with further advancement and evolution of technologies, algorithms and investing strategies various other models could be developed that give more efficient results.

References

[1] Adebiyi A. and Adewumi O. (2014). *Stock Price Prediction Using the ARIMA Model*, IEEE Computer Society.

[2] Mokta Rani Sarker. (2015). *Comparison among Different Models in Determining Optimal Portfolio: Evidence from Dhaka Stock Exchange in Bangladesh,* Journal of Business Studies.

[3] https://www.investopedia.com/

[4] Vinay Kumar. (2018). *A simplified perspective of the markowitz portfolio theory,*IJRARVOLUME 5 ISSUE 3, JULY – SEPT 2018

[5] Ankit Dangi. (2012). *Financial Portfolio Optimization: Computationally guided agents to investigate, analyse and invest!*