Predicting Market Movement using Large Language Models

Jerick Shi (junkais@andrew.cmu.edu)

Advisor: Burton Hollifield

Introduction

Predicting the movement of the stock market and other assets has been valuable over the past few decades. Knowing how the value of a certain sector market may move in the future provides much information for investors, as they use that information to develop strategies in order to maximize profit or minimize risk.

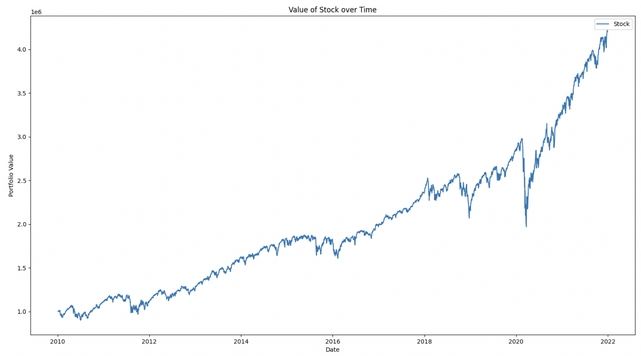

Figure 1. Movement of Stocks over Time

Our goal is to use LLM's such as the GPT model to create a portfolio of assets that minimizes risk. In particular, we hope to create a strategy that outperforms the well-known 60/40 (60% in stocks, 40% in bonds) portfolio. While other research solved similar problems using stock returns, we shall using the correlation coefficient since it is much more plausible to measure and provides benefit to diversifying a portoflio.

Challenges

- **Market Data is Noisy:** The market is extremely volatile and reacts quickly to news. This makes predicting the market much more challenging.
- **News Data is Hard to Consolidate:** While news data is highly abundunt, it is difficult to choose which sources of data reflect the current market.
- **Models are Prone to Overfitting** Regardless to whatever model used, it is highly likely to overfit market data due to the amount of noise it had, making it hard to generalize to new data.

Data Source

The data that we will use will come from the **Beige Book** from the Federal Reserve, which reports the economic conditions in different states in the US 8 times a years. Since this dataset is an official report from the Federal Reserve, the text would not only contain direct information regarding the market, it is also clean data since it is concise.

The 60/40 Portfolio

For several decades, the strategy of putting 60% of your assets in stocks and 40% of your assets has good performance, since it balances the contrasting variances between these 2 assets as well as their negative correlation.

However, after Covid-19, investors had much debate on whether the 60/40 strategy still functions, due to drastic changes in interest rates as well as increased correlation between stocks and bonds. Hence, the 60/40 portfolio is a good baseline strategy to compare against.

Minimizing Variance of a Portfolio

Our goal is to choose a $w \in [0, 1]$ such that putting w of our portfolio in stocks and $1 - w$ of our portfolio in bonds minimizes variance. Mathematically,

$$w^* = argmin_w \, Var(wS + (1 - w)B)$$

in which we can solve for the optimal w^* as follows:

$$w^* = \frac{\sigma_B^2 - \rho\sigma_S\sigma_B}{\sigma_S^2 + \sigma_B^2 - 2\rho\sigma_S\sigma_B}$$

where σ_S and σ_B represent the standard deviations of stocks and bonds respectively, and ρ corresponds to their correlation. We shall use the rolling variances for σ_S and σ_B, and use the GPT model to predict ρ.

Prompt Engineering

We demonstrate our refined prompt below:

```
The following data is an article regarding the economic conditions: [data]

Only using the above data and the fact that previous 3 month's
correlations are given by [correlations] from oldest to most recent, and
nothing else, do you think bonds returns and stock returns will be
positively or negatively correlated?

Do not use any other outside information. You must choose a side, you
don't need to be certain about it. Respond with only 1 or 0, 1 being
positively correlated, 0 being negatively correlated. Then give a number
between 1 and 5 on how confident you are, 1 being least confident, 5
being fully confident. You should return 2 digits separated by a comma
and nothing else.

Example: 1,3. This example implies the guess is positively correlated with
mediocre confidence. Ensure the formatting is correct.
```

Figure 2. Prompt to GPT model to obtain correlation

Simulation

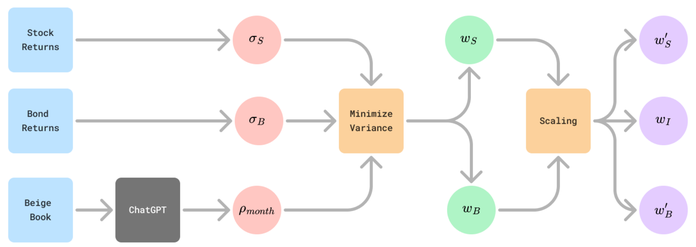

Figure 3. Algorithm for Simulating Portfolio Value

The above algorithm shows how the portfolio value is simulated each month over time, where w_B, w_S, and w_I are the weights in bonds, stocks, and the bank respectively. We add the bank as a risk-free asset in order to fully demonstrate the difference in performances of different strategies under the same variances.

Results/Conclusions

We measure the results of our strategies using the sharpe ratio, which gives a score based off of the returns and stability of the portfolio:

$$\theta_a = \frac{E[R_a - R_f]}{\sigma_a}$$

Strategy	SR (Total)	SR (Before 2022)	SR (2022-2024)
Bond	1.42	2.79	-1.03
60/40	2.68	3.39	**0.48**
GPT	**2.74**	**4.31**	-0.42

Table 1. Sharpe Ratios of Strategies in Different Time Periods

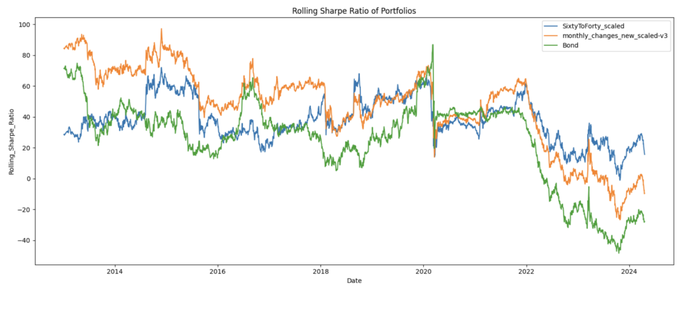

Figure 4. Rolling 3-year Sharpe Ratios of Strategies

From the above data, we make the following conclusions:

- **The Beige Book contains information about Correlations amongst Assets:** We see that during the years before Covid, the GPT strategy outperforms the 60/40, which purely used data from the Beige Book, which indicates that there exists useful information in the Beige Book regarding correlations.
- **The 60/40 is still a good strategy:** We observe that after Covid, the sharpe ratio of the 60/40 is greater than the GPT strategy, both of which is better than putting it all in bonds. Hence, there is no direct evidence that the 60/40 portfolio is outdated.

Future Directions

- **Predict Volatility Movement:** Our current model only predicts correlations, but since the optimal weight also takes in volatility, we can use the GPT model to predict changes in volatility as well.
- **Using other Data Sources:** While the Beige Book is a great resource, we could also incorporate news sources to obtain better results.
- **Using Additional Large Language Models:** With the current race of different LLM's from different companies, other LLM's might outperform the GPT model.
- **Having a more Diverse Portfolio:** The current strategy only allocates the portfolio to the stock and bond index. With an optimized strategy, we can determine how it performs when choosing between multiple assets.