

# Can ChatGPT Improve Investment Decision? From a portfolio management perspective

Hyungjin Ko<sup>a</sup>, Jaewook Lee<sup>a,\*</sup>

<sup>a</sup>*Department of Industrial Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea*

---

## Abstract

ChatGPT as a large language model (LLM) has recently gained significant attention for generating human-like text. While much of the existing research has focused on its applications for writing and language translation, the potential of ChatGPT in finance, particularly in the context of investment, remains unexplored. This research aims to examine the efficacy of ChatGPT in aiding portfolio management with regard to asset allocation and diversification through the use of a quantitative methodology. We use ChatGPT to select asset classes and evaluate the diversification effect of its selections. Our results suggest that ChatGPT's selections are statistically significantly better in terms of diversity index than randomly selected assets. We also construct portfolios based on ChatGPT's selections and find that they outperform portfolios built on randomly selected assets. Overall, our study contributes to a better understanding of the role of LLMs like ChatGPT as potential assistants for portfolio managers.

*Keywords:* ChatGPT, Mean-variance framework, Asset allocation, Portfolio management, Diversification

---

---

\*Corresponding author

*Email addresses:* [hyungjinko@snu.ac.kr](mailto:hyungjinko@snu.ac.kr) (Hyungjin Ko), [jaewook@snu.ac.kr](mailto:jaewook@snu.ac.kr) (Jaewook Lee)

# Can ChatGPT Improve Investment Decision? From a portfolio management perspective

Anonymous<sup>a</sup>

<sup>a</sup>Omitted

---

## Abstract

ChatGPT as a large language model (LLM) has recently gained significant attention for generating human-like text. While much of the existing research has focused on its applications for writing and language translation, the potential of ChatGPT in finance, particularly in the context of investment, remains unexplored. This research aims to examine the efficacy of ChatGPT in aiding portfolio management with regard to asset allocation and diversification through the use of a quantitative methodology. We use ChatGPT to select asset classes and evaluate the diversification effect of its selections. Our results suggest that ChatGPT's selections are statistically significantly better in terms of diversity index than randomly selected assets. We also construct portfolios based on ChatGPT's selections and find that they outperform portfolios built on randomly selected assets. Overall, our study contributes to a better understanding of the role of LLMs like ChatGPT as potential assistants for portfolio managers.

*Keywords:* ChatGPT, Mean-variance framework, Asset allocation, Portfolio management, Diversification

---

## 1. Introduction

Recently, there has been a significant increase in interest in LLM, particularly ChatGPT, a fine-tuned version of the Generative Pre-trained Transformer (GPT) model designed for chatbot applications<sup>1</sup>. With access to a vast dataset, large parameters, and a sizeable terabyte of text data, the software can respond to queries and produce answers that resemble human-like responses based on user input<sup>2</sup>. ChatGPT has become a sensation in the industry, with over one million users signing up in its first week of launch and 616 million monthly website visitors as of our writing (Taecharungroj, 2023). The program reached 100

---

<sup>1</sup><https://openai.com/blog/chatgpt>

<sup>2</sup>The model was trained using a combination of reinforcement learning algorithms and human input, utilizing over 150 billion parameters, as reported by Dowling & Lucey (2023). The training data set was halted in 2021, meaning that the model can only access data from years prior to 2021.

million users by January 2023 and is poised to become a significant disruptor in the industry (Grant & Metz, 2022).

While some studies have examined the role of ChatGPT in writing and language generation (Liebrenz et al., 2023; Thorp, 2023; van Dis et al., 2023; Biswas, 2023; Patel & Lam, 2023; Kitamura, 2023; Chen, 2023; D'Amico et al., 2022; Macdonald et al., 2023; Perkins, 2023; Huh, 2023; Manohar & Prasad, 2023), our study focuses on the potential of ChatGPT as an assistant for portfolio managers in terms of asset allocation regarding diversification effect. Our motivation stems from the idea that, as an LLM, ChatGPT has the potential to capture special aspects of the market that human investors cannot, rendering it a valuable tool for managing portfolios. However, despite its promise, there is a lack of research on the role of ChatGPT in finance, particularly in the area of investment.

Our study aims to investigate ChatGPT's capability to assist investors in selecting assets and achieving portfolio diversification<sup>3</sup>. Some individuals assumed ChatGPT could predict stock prices; however, it is important to note that ChatGPT is not a prediction model. Thus, predicting stock prices is beyond its capabilities<sup>4</sup>. This leads to the question of whether ChatGPT can be used for investing. To tackle this issue, we reframed the problem by assuming the role of an assistant to a portfolio manager, asking different types of questions rather than expecting ChatGPT to predict the future. We obtain ChatGPT's asset selection by questioning the model and compute three diversity indices to evaluate its selections. Additionally, we compare the performance of portfolios constructed based on ChatGPT's selections to portfolios built on randomly selected assets.

Based on our findings, ChatGPT's selections display a statistically significant improvement in diversity index compared to randomly selected assets, demonstrating its asset selection process is based on an asset allocation philosophy emphasizing diversification. Additionally, our analysis of portfolio performance indicates that portfolios constructed using ChatGPT's selections outperform those constructed using randomly selected assets, indicating that ChatGPT's selections have a sufficient diversification effect at the portfolio level, thereby enabling users to build more efficient portfolios. Overall, ChatGPT can identify abstract relationships between assets, particularly in terms of their dissimilarity with respect to the asset class.

Dowling & Lucey (2023) studied ChatGPT's usefulness for financial writing. However, no research has explored ChatGPT's role in investment in the finance literature. Our study contributes to this literature by examining the potential of LLM in the financial field and expanding our understanding of its role beyond language generation and writing, specifically in asset allocation and portfolio

---

<sup>3</sup>In portfolio management, the most crucial factors that impact the ultimate performance of the portfolio are the selection of asset classes and the corresponding weightings assigned to each of them (Brinson et al., 1986).

<sup>4</sup>Even human investors find it challenging or impossible to predict the future accurately, making it unsurprising that LLM cannot either.

management. Our research provides timely insights into the versatility of ChatGPT as an assistant for portfolio management. We discovered that ChatGPT could make appropriate asset selections that would diversify the final portfolio and thus reveal that ChatGPT could potentially serve as a valuable co-pilot for portfolio management in the finance industry.

The structure of our study is as follows. In Section 2, we detail the methodology employed. Section 3 outlines our data and experimental design. In Section 4, we present our empirical findings. Finally, Section 5 concludes our study.

## 2. Methodology

### 2.1. Diversity indices

#### 2.1.1. The Simpson and Shannon diversity indices

We employ Simpson and Shannon diversity indices to evaluate the level of diversity in the assets that ChatGPT selects. These metrics are suited to assessing our scenario, in which a fixed number of assets are chosen to construct a portfolio. It is desirable for selected assets to originate from a range of diverse asset classes rather than from a single class, as this facilitates greater diversification effects on the final portfolio. The Simpson index is calculated as follows:

$$I_{Simpson} = 1 - \sum_{i=1}^s p_i^2, \quad (1)$$

where  $p_i$  represents the proportion of the total sample that is comprised of the  $i$ -th asset class and  $s$  is the number of classes. The Shannon index is formulated as follows:

$$I_{Shannon} = - \sum_{i=1}^s p_i \ln(p_i). \quad (2)$$

#### 2.1.2. The average correlation

In addition, we utilize the average correlation as a measure of diversification. The correlation between returns on assets can act as an indicator for capturing distinct movements that offer diversification benefits for risk mitigation. To calculate the average correlation, we obtain the upper-right triangular elements of the correlation matrix and add them up, dividing the sum by the number of elements, as shown below:

$$\rho_{avg} = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^k \rho_{ij}, \quad (3)$$

where  $\rho_{ij}$  is the correlation coefficient between the  $i$ -th and  $j$ -th assets,  $k$  is the number of assets, and the factor of 2 in the numerator accounts for the fact that only the upper triangular elements are used.

## 2.2. Portfolio analysis based on the mean-variance framework

To conduct portfolio analysis, we employ the mean-variance framework, which is a widely used methodology proposed by Markowitz (1952). After asset selection, portfolio managers typically construct portfolios as their second core task by solving an optimization problem with an objective function, such as risk or risk-adjusted return. The performance of the resulting portfolio is then evaluated to assess the degree of diversification achieved at the portfolio level through ChatGPT’s decision-making.

The mean-variance framework can be represented as follows:

$$\begin{aligned} \min_w \quad & \frac{\lambda}{2} \mathbf{w}^T \Sigma \mathbf{w} - \mathbf{w}^T \boldsymbol{\mu}, \\ \text{s.t.} \quad & \mathbf{w}^T \mathbf{1} = 1 \\ & w_i \geq 0 \quad (i = 0, \dots, K) \end{aligned} \quad (4)$$

where  $\mathbf{w} \in \mathbb{R}^K$  denotes the weight vector of asset,  $\Sigma \in \mathbb{R}^{K \times K}$  denotes the variance-covariance matrix of asset returns,  $\boldsymbol{\mu} \in \mathbb{R}^K$  denotes the expected return vector, and  $\lambda \in \mathbb{R}^1$  denotes the risk aversion coefficient of investors. We employ the mean-variance framework to generate two distinct portfolio types: a tangency portfolio<sup>5</sup> and a minimum risk portfolio<sup>6</sup>.

In summary, we utilize diversity indices and average correlation to evaluate ChatGPT’s capability to select diverse assets and the mean-variance framework to assess the degree of diversification achieved at the portfolio level through ChatGPT’s decision-making. Through these, we can elucidate the potential role of ChatGPT in the financial field and its ability to serve as an investment assistant for portfolio managers.

## 3. Data and experimental design

### 3.1. Data

We collected a data sample of 20 large-cap stocks from various sectors<sup>7</sup> listed on the US stock market, along with five assets from each of the following classes: cryptocurrencies, commodities, currencies, and bonds, resulting in a total of 40 assets<sup>8</sup>. The data sample covers the period from January 1, 2021, to January 31, 2023<sup>9</sup>. We calculated simple returns on a daily basis<sup>10</sup> and used the 3-month

---

<sup>5</sup>The tangency portfolio is derived by maximizing the Sharpe ratio (SR) (equivalent to solving Eq. (4) with  $\boldsymbol{\mu} \neq 0$ ), which gauges the degree of diversification achieved at the portfolio level.

<sup>6</sup>The minimum risk portfolio is derived by minimizing the variance of the portfolio (equivalent to solving Eq. (4) with  $\boldsymbol{\mu} = 0$ ), which measures the level of risk.

<sup>7</sup>These include IT, industrials, health care, real estate, and finance sectors.

<sup>8</sup>Our selection criteria were based on market capitalization, popularity, and representativity for investors.

<sup>9</sup>All prices are obtained from [www.finance.yahoo.com](http://www.finance.yahoo.com) and are denominated in US dollars.

<sup>10</sup>The in-sample period consists of 252 observations, while the out-of-sample period has 274 observations.

Table 1: The summary statistics for the returns of our data

	$N$	Mean	Median	Min	Max	Stdev.	Skewness	Kurtosis
Stock	20	0.120	0.096	-42.3	11.9	0.299	0.122	2.59
Cryptocurrency	5	0.925	0.374	-83.8	53.2	1.044	1.198	13.1
Commodity	5	0.003	-0.064	-36.2	7.70	0.288	0.010	1.44
Currency	5	-0.029	-0.030	-10.4	4.00	0.081	0.248	3.89
Bond	5	-0.031	-0.096	-10.0	4.10	0.123	0.207	2.81
Total	40	0.169	0.036	-83.8	4.00	0.341	0.269	3.95

*Notes.* Table 1 presents summary statistics for the returns of our data, including the number of observations, mean, median, minimum, maximum, standard deviation, skewness, and kurtosis. The table shows the statistics for five different asset classes: stocks, cryptocurrencies, commodities, currencies, and bonds. For the sake of concision, we have included detailed asset names and tickers in Table A1 in the appendix.

treasury-bill rate as a proxy for the risk-free rate. Table 1 provides summary statistics for the returns of our data.

### 3.2. Experimental design

For portfolio managers or investors, asset selection is a crucial undertaking, as the performance of a portfolio is heavily reliant on a suitable mix of diverse asset classes (Brinson et al., 1986). Nevertheless, identifying and differentiating between assets from different classes is challenging, particularly for those with limited knowledge of portfolio choice and asset allocation theory. As such, we analyze the outcomes of ChatGPT’s decision-making by evaluating its performance in selecting diversified assets and comparing it to random selection as a baseline.

To evaluate ChatGPT’s effectiveness in asset selection, we carried out a series of experiments. We initially requested ChatGPT to select  $k$  assets from 40 assets across five asset classes. We repeated this process by randomly selecting  $k$  assets. Diversity indices were computed, and tangency and minimum risk portfolios were constructed using assets from the previous steps. We assessed the portfolios’ SRs and risks for out-of-sample data. To ensure reliability and confidence in our results, we repeated this process 10,000 times.

We used 1-year<sup>11</sup> return data from January 1, 2021, to December 31, 2021, for the in-sample period to estimate the expected return and covariance<sup>12</sup>. To avoid instability in constructing the portfolio with a larger number of assets due to the small number of observations (252), we set  $k = 2, 3, 4$ <sup>13</sup>.

<sup>11</sup>The common practice of simply averaging returns by incorporating long-term historical data is often regarded as an inadequate and unstable approach to estimating portfolio returns. To ensure stability and avoid the potential negative impact of distant past data, we employ a one-year time horizon, which provides information that is reasonably moderate in length.

<sup>12</sup>We did not rebalance the portfolio due to the short length of the sample, instead maintaining the estimated weight to backtest the entire out-of-sample period.

<sup>13</sup>We followed the practice of DeMiguel et al. (2009) and set the risk aversion coefficient  $\lambda = 1$ .

ChatGPT is a language model trained on data up to December 31, 2021<sup>14</sup>. We consider a total of 13 months from January 1, 2022, to January 31, 2023, as the out-of-sample period<sup>15</sup>. We acknowledge that our study is limited by the short amount of available data, but we conducted various empirical experiments to ensure the reliability of our findings.

In summary, we have selected an appropriate number of assets across multiple classes to ensure tractability and included a variety of asset classes to effectively evaluate ChatGPT’s differentiation ability. Our design is comprehensive and reliable for portfolio analysis.

## 4. Empirical Findings

### 4.1. Diversity indices results

Table 2: The results of three diversity indices

	Random	ChatGPT	t-value
Panel A: the Simpson Diversity Index			
N=2	0.334	<b>0.380</b>	-4.574***
N=3	0.340	<b>0.542</b>	-22.36***
N=4	0.344	<b>0.548</b>	-30.33***
Panel B: the Shannon Diversity Index			
N=2	0.668	<b>0.760</b>	-4.574***
N=3	0.680	<b>1.233</b>	-27.54***
N=4	0.689	<b>1.316</b>	-30.33***
Panel C: Average correlation			
N=2	0.186	<b>0.170</b>	1.707*
N=3	0.172	<b>0.141</b>	4.412***
N=4	0.159	<b>0.144</b>	2.567**

*Notes.* Table 2 presents the results of diversity measures of asset selection by ChatGPT compared to random selection. Panel A and B show the Simpson and Shannon diversity indices, respectively, for asset selections with two to four assets. The t-value column presents the statistical significance of the difference between ChatGPT and random selection. Panel C presents the average correlation between asset returns for ChatGPT’s asset selection and random selection. The results indicate that ChatGPT’s asset selection is more diverse than random selection based on both diversity indices, and ChatGPT’s asset selection provides a higher potential for diversification effect based on the average correlation measure. All t-tests are significant at the 1% level or higher.

We analyze the level of diversity in ChatGPT’s asset selection compared to a randomly selected group of assets using Simpson and Shannon diversity indices. Our results demonstrate that ChatGPT’s asset selection has higher values for both diversity indices than the random selection, irrespective of the number of assets. This indicates that ChatGPT’s asset selection is more diverse than the randomly selected assets. The t-test results presented in the “t-value” column further support our findings by demonstrating statistical significance.

<sup>14</sup>When asked about certain information, such as the value of the Fed rate, ChatGPT states that the target federal funds rate range set by the Federal Reserve is 0.00%-0.25%, which is not true at the time of writing.

<sup>15</sup>At the time of writing, data up to January 31, 2023, is available.

Additionally, we investigate the average correlation of ChatGPT's asset selection and observe that it is lower than that of the randomly selected assets. This finding indicates that ChatGPT's asset selection has the potential for higher diversification effects, which is further corroborated by the t-test results.

Our discovery regarding diversity indices suggests that ChatGPT's asset selection reflects an asset allocation philosophy in terms of diversification effects rather than random selection.

#### 4.2. Portfolio analysis results

Table 3: The out-of-sample results of the tangency portfolio.

	Expected return	Risk		Risk-adjusted return	
		Stdev.	t-value	SR	t-value
Panel A: k=2					
Random	-0.170	0.409	-	-0.361	-
ChatGPT	-0.117	<b>0.353</b>	7.457***	<b>-0.328</b>	-2.261**
Panel B: k=3					
Random	-0.208	0.470	-	-0.378	-
ChatGPT	-0.115	<b>0.324</b>	18.62***	<b>-0.262</b>	-8.762***
Panel C: k=4					
Random	-0.231	0.508	-	-0.388	-
ChatGPT	-0.133	<b>0.332</b>	22.09***	<b>-0.245</b>	-10.52***

*Notes.* Table 3 presents the out-of-sample results of the tangency portfolio for different asset selections of size  $k = 2, 3, 4$ , comparing the performance of portfolios constructed based on ChatGPT's selection and random selection. The table reports expected return, risk (measured by standard deviation), and risk-adjusted return (measured by SR) along with the corresponding t-values. The results show that, in general, ChatGPT outperforms random selection in terms of risk and risk-adjusted return, with statistically significant t-values. All portfolio values of expected returns are negative, but this outcome is not surprising due to the challenging market conditions during the out-of-sample periods. These conditions should be considered when interpreting the portfolio performance results.

Next, we examine the out-of-sample performance of the tangency and minimum risk portfolios based on ChatGPT's selected assets. We first examine the tangency portfolios, and our results show that portfolios selected using ChatGPT have lower risks compared to random selection-based portfolios, regardless of the number of assets. The statistical significance of this finding is supported by the t-test. However, it is noteworthy that the tangency portfolio is equivalent to the maximum SR portfolio. Therefore, the key focus should be on the improvement in the SR observed across ChatGPT-based portfolios, irrespective of the number of assets. Our results indicate a significant improvement in SR when using ChatGPT-based asset selection compared to random selection. This suggests that ChatGPT can help improve the diversification effect of asset selection, as supported by the t-test statistics.

Interestingly, the SR of ChatGPT portfolios increases as the number of assets included in the portfolio increases, in contrast to random selection-based portfolios. The gap between the SR of ChatGPT and random selection-based portfolios widens as the number of assets increases. These findings suggest that



ChatGPT is able to identify and capture the distinct movements of different asset classes, resulting in a pronounced diversification effect.

Table 4: The out-of-sample results of the minimum risk portfolio.

	Expected return	Risk		Risk-adjusted return	
		Stdev.	t-value	SR	t-value
Panel A: k=2					
Random	-0.076	0.216	-	<b>-0.407</b>	-
ChatGPT	-0.046	<b>0.165</b>	12.71***	-0.434	1.551
Panel B: k=3					
Random	-0.069	0.171	-	-0.487	-
ChatGPT	-0.007	<b>0.119</b>	16.88***	<b>-0.16</b>	-17.68***
Panel C: k=4					
Random	-0.063	0.140	-	-0.560	-
ChatGPT	0.005	<b>0.112</b>	11.18***	<b>-0.054</b>	-27.15***

*Notes.* This table shows the out-of-sample results of the minimum risk portfolio, with three panels for different numbers of assets ( $k$ ). Each panel includes two rows, one for the randomly selected assets and the other for assets selected by ChatGPT. The columns show the expected return, risk, and risk-adjusted return measured by the SR, with the corresponding t-values indicating statistical significance. The results demonstrate that ChatGPT-based minimum risk portfolios outperform random selection in terms of risk and risk-adjusted returns with the exception of  $k = 3$  in SR. Specifically, ChatGPT's portfolios have lower risk and higher risk-adjusted returns, as indicated by the lower standard deviation and higher SR values, respectively.

Regarding minimum risk portfolios, ChatGPT-based portfolios exhibit lower risks than those of randomly selected assets, demonstrating that ChatGPT can significantly enhance the overall risk management of the portfolio. The t-test results support this finding. When  $k = 3, 4$ , the SRs of ChatGPT-based portfolios are better than those of random selection. However, when  $k = 2$ , the SR of ChatGPT-based portfolios is not better than random selection because the minimum risk portfolio aims to minimize portfolio risk, not maximize the SR. Furthermore, it is not surprising that since all portfolio backtesting is conducted out-of-sample, the estimated weights do not always ensure improvements or meet the primary objective in portfolio-level performance due to in-sample overfitting.

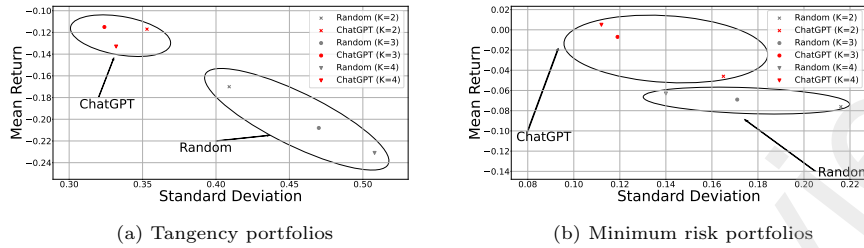


Figure 1: This figure illustrates the risk-return perspective of tangency portfolios (Panel (a)) and minimum risk portfolios (Panel (b)). The expected returns and risks of all portfolios are calculated by averaging all 10,000 portfolios' returns and risks for all scenarios. The red color represents ChatGPT-based portfolios, and the grey represents random selection-based portfolios. The X-, O-, and triangle-shaped markers denote  $k = 2, 3,$  and  $4,$  respectively.

The outcomes presented in the preceding tables are in line with the risk-return profile plot in Figure 1. This plot showcases the risk-return perspective of the tangency (Panel (a)) and minimum risk (Panel (b)) portfolios. To calculate the expected returns and risks of all portfolios, the return and risk of 10,000 portfolios were averaged for all scenarios. In Panel (a), the results of ChatGPT are grouped together in the upper-left quadrant, while those of random selection are clustered in the southeast. This indicates that ChatGPT's tangency portfolios are more efficient than random selection's tangency portfolios, demonstrating that ChatGPT-based portfolios offer a diversification effect on overall portfolios and outperform those selected randomly. In Panel (b), two clusters are evident: ChatGPT and random selection-based minimum risk portfolios. On average, ChatGPT-based minimum risk portfolios exhibit lower risks than those of random selection, implying the efficiency of ChatGPT-based portfolios. Moreover, ChatGPT-based portfolios outperform random selection in the given  $k$  values. In conclusion, ChatGPT's selection offers a diversification effect and achieves effective risk management in portfolio construction.

## 5. Conclusion

This study aimed to investigate the effectiveness of the LLM, ChatGPT, as a tool for portfolio management. Our findings suggest that ChatGPT can be a helpful assistant in the investment process, providing plausible and credible explanations<sup>16</sup> for a given set of asset classes. Although ChatGPT cannot predict the future or optimize portfolios to the same extent as human portfolio managers, it does offer a convenient and practical option for individuals lacking knowledge or proficiency in portfolio management.

From an investment perspective, our study highlights the potential benefits of utilizing ChatGPT as an aid in managing portfolios. Retail investors, particularly those who may be uneducated or misinformed, can benefit from the

<sup>16</sup>A detailed explanation of ChatGPT's decision-making process is presented in Table A2.

democratization of portfolio management. Additionally, professional portfolio managers can improve their productivity by focusing on more important tasks while being assisted by ChatGPT in the selection of diverse assets for a given portfolio.

ChatGPT can also be employed as an educational resource for inexperienced investors or individuals who lack knowledge of financial matters. ChatGPT can impart fundamental financial concepts<sup>17</sup> in a comprehensible and user-friendly manner, providing valuable guidance and supportive education in the field of finance<sup>18</sup>.

Overall, ChatGPT represents an important advancement in the democratization and accessibility of portfolio management and financial education. While our study does have limitations, particularly in terms of short-term validation, we believe that the findings suggest promising avenues for future research. Specifically, we recommend further investigation into the potential for ChatGPT to improve asset pricing determinism and the long-term validation of its use in portfolio management.

In conclusion, ChatGPT is not a “prophet” of the financial industry but rather an “assistant” or “co-pilot” to investors and portfolio managers. When used appropriately within a set framework, ChatGPT has the potential to revolutionize asset allocation practices. However, it is important to note that the role of human experts in portfolio management remains crucial, as they must work in tandem with theoretical frameworks and investment theory to ensure proficiency and reliability in verifying accurate information from LLM.

## References

- Arif, T. B., Munaf, U., & Ul-Haque, I. (2023). The future of medical education and research: Is chatgpt a blessing or blight in disguise?
- Biswas, S. (2023). Chatgpt and the future of medical writing.
- Brinson, G. P., Hood, L. R., & Beebower, G. L. (1986). Determinants of portfolio performance. *Financial Analysts Journal*, 42, 39–44.
- Chen, T.-J. (2023). Chatgpt and other artificial intelligence applications speed up scientific writing. *Journal of the Chinese Medical Association*, (pp. 10–1097).
- D’Amico, R. S., White, T. G., Shah, H. A., & Langer, D. J. (2022). I asked a chatgpt to write an editorial about how we can incorporate chatbots into neurosurgical research and patient care. . . .

---

<sup>17</sup>Please refer to Table A2.

<sup>18</sup>There have been several recent studies exploring the potential educational benefits of ChatGPT. Many of these studies have focused on general and medical education contexts, demonstrating the potential for ChatGPT to serve as a useful educational tool in these fields (Kung et al., 2023; Gilson et al., 2023; Arif et al., 2023; Tlili et al., 2023).

- DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies*, 22, 1915–1953.
- van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). Chatgpt: five priorities for research. *Nature*, 614, 224–226.
- Dowling, M., & Lucey, B. (2023). Chatgpt for (finance) research: The bananarama conjecture. *Finance Research Letters*, (p. 103662).
- Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., Chartash, D. et al. (2023). How does chatgpt perform on the united states medical licensing examination? the implications of large language models for medical education and knowledge assessment. *JMIR Medical Education*, 9, e45312.
- Grant, N., & Metz, C. (2022). A new chat bot is a ‘code red’ for google’s search business.
- Huh, S. (2023). Emergence of the metaverse and chatgpt in journal publishing after the covid-19 pandemic. *Science Editing*, 10, 1–4.
- Kitamura, F. C. (2023). Chatgpt is shaping the future of medical writing but still requires human judgment.
- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J. et al. (2023). Performance of chatgpt on usmle: Potential for ai-assisted medical education using large language models. *PLOS Digital Health*, 2, e0000198.
- Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with chatgpt: ethical challenges for medical publishing. *The Lancet Digital Health*, 5, E105–E1063.
- Macdonald, C., Adeloye, D., Sheikh, A., & Rudan, I. (2023). Can chatgpt draft a research article? an example of population-level vaccine effectiveness analysis. *Journal of Global Health*, 13.
- Manohar, N., & Prasad, S. S. (2023). Use of chatgpt in academic publishing: A rare case of seronegative systemic lupus erythematosus in a patient with hiv infection. *Cureus Journal of Medical Science*, 15.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7, 77–91.
- Patel, S. B., & Lam, K. (2023). Chatgpt: the future of discharge summaries? *The Lancet Digital Health*, 5, E107–E108.
- Perkins, M. (2023). Academic integrity considerations of ai large language models in the post-pandemic era: Chatgpt and beyond. *Journal of University Teaching & Learning Practice*, 20, 07.

Taecharungroj, V. (2023). "what can chatgpt do?" analyzing early reactions to the innovative ai chatbot on twitter. *Big Data and Cognitive Computing*, 7, 35.

Thorp, H. H. (2023). Chatgpt is fun, but not an author.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: Chatgpt as a case study of using chatbots in education. *Smart Learning Environments*, 10, 1-24.

## Appendix

Table A1: The detailed names and tickers of the assets in our data.

Asset	Number	Name (Ticker)
Stock	20	Apple Inc. (AAPL), Microsoft Corp. (MSFT), NVIDIA Corp. (NVDA), Salesforce, Inc. (CRM), Adobe Inc. (ADBE), UnitedHealth Group Inc. (UNH), United Parcel Service, Inc. (UPS), Union Pacific Corp. (UNP), Johnson & Johnson (JNJ), Eli Lilly and Company (LLY), Prologis, Inc. (PLD), American Tower Corp. (AMT), Equinix, Inc. (EQIX), Crown Castle Inc. (CCI), Public Storage (PSA), JPMorgan Chase & Co. (JPM), Visa Inc. (V), Bank of America Corp. (BAC), Mastercard Inc. (MA), and Morgan Stanley (MS).
Cryptocurrency	5	Bitcoin (BTC-USD), Ethereum (ETH-USD), Binance coin (BNB-USD), Cardano (ADA-USD), and Ripple (XRP-USD).
Commodity	5	SPDR Gold Shares (GLD), iShares Silver Trust (SLV), abrdn Physical Palladium Shares ETF (PALL), abrdn Physical Platinum Shares ETF (PPLT), and United States Copper Index Fund, LP (CPER).
Currency	5	Invesco DB US Dollar Index Bullish Fund (UUP), USD/EUR (EURUSD=X), USD/GBP (GBPUSD=X), USD/JPY (JPYUSD=X), and USD/CNY (CNYUSD=X).
Bond	5	iShares U.S. Treasury Bond ETF (GOVT), Vanguard EUR Eurozone Government Bond UCITS ETF (VETAL), Xtrackers II Harvest China Government Bond UCITS ETF (OGB.MI), iShares Core UK Gilts UCITS ETF GBP (IGLTL), and Xtrackers II Japan Government Bond UCITS ETF (XJSE.MI).

*Notes.* This table shows the asset classes included in the study, along with their specific names, tickers, and the number of assets. The asset classes are divided into five categories: stock, cryptocurrency, commodity, currency, and bond. The stock category includes 20 assets, while the other categories each include five assets. The tickers for each asset are listed in parentheses after their names.

On what basis does ChatGPT make its selections? To address this inquiry, we provide an elaborate description of ChatGPT’s decision-making process in Table A2. Our investigation into the rationale behind ChatGPT’s selections reveals that they incorporate key principles of portfolio management, including diversification, hedge, and safe haven effects, sector and industry diversity, knowledge of each industry, downside risk, and correlation. These findings suggest that ChatGPT’s selections are based on established principles of portfolio management, lending support to the validity of its selections.

In summary, our results indicate that ChatGPT’s outputs are appropriate and plausible recommendations that can enhance diversification at the asset selection and portfolio levels with a convincing narrative.

Table A2: Assets selected by ChatGPT with detailed explanations of the reasons behind each selection.

Selected assets	Reasons
<p>Example 1: JPM (JPMorgan Chase &amp; Co.), GLD (SPDR Gold Shares), and UNH (UnitedHealth Group).</p>	<p>1. JPM: JPM is a global financial services firm that offers a range of financial products and services, including investment banking, asset management, and consumer banking. Including JPM in the portfolio would offer a degree of financial sector exposure, thereby enhancing <b>diversification</b>. This is because the financial sector often <b>behaves differently from other sectors</b>, making it a favorable choice for diversification. Additionally, JPM has a history of strong financial performance and a reputation for being a well-managed company.</p> <p>2. GLD: This exchange-traded fund (ETF) tracks the performance of gold bullion, which has historically been considered a <b>safe haven</b> asset. As a result, GLD can act as a <b>hedge</b> against inflation and <b>economic uncertainty</b>. Gold has a <b>low correlation</b> with other asset classes, which makes it an effective <b>diversification</b> tool.</p> <p>3. UNH: This healthcare company is the largest health insurer in the US, and it has a diversified business model that includes healthcare services, technology, and pharmacy benefits management. UNH has a strong financial position and has consistently increased its dividend over the years. Additionally, the healthcare industry is relatively immune to <b>economic cycles</b>, which makes UNH an attractive investment option for investors looking to <b>minimize risk</b>.</p>
<p>Example 2: MSFT (Microsoft Corp.), GOVT (iShares U.S. Treasury Bond ETF), and SLV (iShares Silver Trust).</p>	<p>1. MSFT: Microsoft is a technology company that operates across various segments, including productivity and business processes, intelligent cloud, and personal computing. This provides a <b>diversification</b> opportunity in a <b>different sector</b>. Microsoft is a highly profitable company with a strong balance sheet.</p> <p>2. GOVT: GOVT is the iShares U.S. Treasury Bond ETF, which seeks to track the investment results of an index composed of U.S. Treasury bonds. Investing in U.S. Treasury bonds can provide <b>diversification benefits</b> to your portfolio, as they are generally considered to be <b>safe-haven</b> assets that can help <b>reduce overall portfolio risk</b>. Additionally, bonds typically have a <b>lower correlation</b> with stocks and other asset classes, which can help reduce the overall volatility of a portfolio.</p> <p>3. SLV: Silver has a <b>low correlation</b> with other asset classes like stocks and bonds. Additionally, SLV provides exposure to the performance of the price of silver, which has historically been a good <b>diversifier</b> in a portfolio.</p>

*Notes.* This table offers a selection of recommended assets accompanied by a detailed explanation of the rationale behind each selection. The table features two columns: one that lists the selected assets and the other that outlines the reasons for their inclusion. To ensure conciseness, we only provide illustrations for the scenario where  $k = 3$ . Notably, ChatGPT appears to have a good grasp of various financial concepts, including asset allocation, portfolio management, risk management, and portfolio choice. The selections in the table reflect this understanding, as they take into account key factors such as the diversification effect, hedging, safe haven assets, correlation, economic uncertainty, economic cycles, risk, and sector exposure, among others, as denoted in bold text.