Enhancing Portfolio Optimization with Transformer-GAN Integration: A Novel Approach in the Black-Litterman Framework

1st Enmin Zhu
Computer and Information Science
FST
Macau, China
mc35330@umac.mo

Abstract—This study presents an innovative approach to portfolio optimization by integrating Transformer models with Generative Adversarial Networks (GANs) within the Black-Litterman (BL) framework. Capitalizing on Transformers' ability to discern long-range dependencies and GANs' proficiency in generating accurate predictive models, our method enhances the generation of refined predictive views for BL portfolio allocations. This fusion of our model with BL's structured method for merging objective views with market equilibrium offers a potent tool for modern portfolio management, outperforming traditional forecasting methods. Our integrated approach not only demonstrates the potential to improve investment decision-making but also contributes a new approach to capture the complexities of financial markets for robust portfolio optimization.

Index Terms—Transformer, GAN, Black and Litterman model, Portfolio optimization

I. Introduction

The Black-Litterman model (BL), a renowned tool in modern portfolio theory, offers a structured method for combining investor views with market equilibrium, thereby enabling more tailored and robust portfolio allocations. It significantly mitigates the limitations inherent in the traditional Markowitz mean-variance optimization, especially those related to input sensitivity and estimation errors [1] [2]. However, generating subjective views, a critical component of this model, remains a challenge.

With the advancement of artificial intelligence techniques, scholars have explored various methodologies to generate investor views for the Black-Litterman (BL) model. CNNs, by leveraging multiple indicators, and LSTMs, through historical price data analysis, have significantly contributed to enhancing portfolio performance within the BL framework [3] [4] [5]. However, the complexity of processing numerous indicators increases computational demands, while LSTMs face challenges in capturing long-range dependencies [6]. This necessitates the exploration of new models that can address these limitations and improve price prediction accuracy.

Originally designed for language translation, Transformers have demonstrated exceptional versatility across domains such as computer vision [7], speech recognition [8], and Large Language Models (LLMs) [6]. Their unique architecture, featuring attention mechanisms, enables effective sequential data management, surpassing traditional recurrent and convolutional network models. In the realm of stock price prediction, studies have indicated that Transformer-based models outperform their CNN and LSTM counterparts [9] [10].

Furthermore, Generative Adversarial Networks (GANs) [11] which are constructed by generator and discriminator, have proven adept at uncovering complex patterns in historical stock data, a task challenging for simpler models. Research demonstrates that GAN-based methodologies excel in time series forecasting and stock prediction, offering innovative performance improvements [12] [13] [14] [15].

In the pursuit of enhancing the Black-Litterman model's capabilities, we have innovated an architecture that synergizes the generative power of GANs with the intricate data processing prowess of Transformers. The resulting framework is aimed at capturing the intricacies and potential patterns within financial time series data. Here are the salient features of our approach:

- **Advanced Architectural Integration:** Our model innovatively combines the generative capabilities of GANs with the deep learning prowess of Transformers, adept at uncovering long-range dependencies and latent patterns within historical stock price data.
- **Performance and Adaptability:** The proposed model demonstrates competitive performance metrics, establishing new frontiers in accuracy and scalability, which are corroborated by rigorous benchmarking against traditional GANs and showcasing exceptional adaptability across various data resolutions.
- **Strategic Asset Allocation:** The application of our integrated model to the Black-Litterman framework facilitates a more nuanced and adaptive approach to portfolio construction. By accurately capturing market trends and investor sentiments, it guides asset allocation decisions that can potentially outperform standard market strategies.

II. RELATED WORK

The application of the Black-Litterman (BL) model is contingent upon the precise estimation of parameters derived from return series data. This section focuses on exploring mechanisms that bolster the efficacy of historical data in enhancing portfolio performance. Thus, we scrutinize cutting-edge deep learning algorithms renowned for their predictive prowess and exceptional capability in capturing latent patterns and salient features within financial datasets.

Convolutional Neural Networks (CNNs) utilize convolutional kernels to autonomously extract pivotal features from stock price charts or technical indicators, thereby achieving remarkable accuracy [16] [17]. Yet, CNNs do not inherently account for the sequential nature of stock prices, necessitating additional transformations of time-series data [18]. Addressing this, Long Short-Term Memory networks (LSTM), a specialized subset of Recurrent Neural Networks (RNN), employ 'forget gate' and 'input gate' mechanisms to modulate information flow, significantly boosting their efficacy in processing time-series data in stock analysis [16] [19]. However, despite their prowess in temporal feature capture, LSTMs are less adept at extracting features from high-dimensional spatial data. To mitigate the shortcomings of singular models in predicting stock prices, the CNN-LSTM hybrid model has been proposed, merging CNNs' robust feature extraction capabilities with LSTMs' proficiency in time-series analysis [18] [20] [21]. Further advancements include integrating the attention mechanism into the CNN-LSTM framework, sharpening the model's focus on crucial temporal information and thereby enhancing prediction accuracy and efficiency [22] [23] [24]. Leveraging their proficiency with time series data, models like CNN, LSTM, and their hybrids have been instrumental in forecasting future prices and integrating these predictions as views into the Black-Litterman framework. Barua et al. [5] introduced a CNN-BiLSTM model for predicting index closing prices, subsequently incorporating these forecasts as investor views within the Black-Litterman model. Similarly, Rezaei et al. [4] demonstrated that the hybrid CEEMD-CNN-LSTM model achieves outstanding performance in view generation, highlighting the potential of combining advanced machine learning techniques for financial forecasting and portfolio optimization.

The attention mechanism [25], pivotal to the transformer model, has revolutionized sequential data analysis, enhancing models' focus on pertinent input segments for more effective processing, particularly with lengthy sequences [25]. The transformer distinguishes itself by processing data in parallel, its scalability, and handling long-range dependencies, significantly improving prediction accuracy and modeling complex patterns. While Chen et al. [26] and Zhang et al. [27] have straightforwardly utilized the Transformer for stock market forecasts, its self-attention mechanism's ability to capture long-term dependencies is particularly notable. This feature has led Wang et al. [9] and Zhang et al. [10] to successfully employ the transformer for stock price predictions, markedly outperforming traditional methods. The model's adaptability to various data lengths and parallel processing capabilities make it a powerful tool for detailed and accurate large-scale data analysis.

Generative Adversarial Networks (GANs) [11] have shown remarkable capability in capturing potential patterns or features from complex datasets for their potential to learn high-dimensional, complex data distribution [28]. For instance, Zhang et al. discuss the use of GANs in recovering hidden cell-lineage developmental knowledge from complex biological datasets, showcasing the model's ability to capture essential data semantics in its latent space [29]. Moreover, GANs' application in anomaly detection further exemplifies their capacity to identify abnormal instances by learning potential patterns within data, as explored by Rani et al. [30]. GANs have shown promise in the domain of time series forecasting. Ferchichi et al. [15] introduced a novel multi-attention GAN model, comprising three main components, which achieves notable accuracy across multiple future time intervals. Yin et al. [14] employed a multi-attention GAN framework for stock market prediction, while Polamuri et al. [13] introduced GAN-HPA, yielding promising results in stock price forecasting. Similarly, study of Diqi et al. [12] on StockGAN showcases its efficacy in financial time series analysis, underscoring the growing utility of GAN-based models in predicting market dynamics.

III. BACKGROUND

In the realm of machine learning, various architectures have been developed to tackle a wide range of tasks, each with unique strengths and applications. This section provides an overview of four pivotal models: Convolutional Neural Networks (CNNs), Long Short-Term Memory networks (LSTMs), Generative Adversarial Networks (GANs), and Transformers. CNNs are renowned for their image processing prowess, LSTMs for their ability to capture time-dependent information, GANs for their generative capabilities in producing new, synthetic instances of data, and Transformers for their advanced handling of sequential data through attention mechanisms. We will discuss the key features, foundational principles, and the specific roles these models play in the context of our research.

A. GAN

The realm of Artificial Intelligence (AI) has seen remarkable advancements, leading to significant enhancements in machines' capabilities to mimic human traits. A central strategy in emulating authentic human behavior is through the development of sophisticated machine learning algorithms. Among these, Generative Adversarial Networks (GANs), introduced by Goodfellow et al. in 2014 [11], stand out for their unique operational framework. GANs are predicated on the concept of a two-player zero-sum game, where the sum of the gains and losses between the players is zero, ensuring that any gain or loss in utility by one player is exactly offset by the loss or gain in the other player's utility.

A GAN consists of two main components: the Generator (G) and the Discriminator (D), which are refined through adver-

sarial learning. The Generator's role is to create new, unique input samples that mimic the structure of real data, starting from random noise, despite having no prior access to the real data itself. Conversely, the Discriminator evaluates these samples, distinguishing between the authentic and generated data, aiming to identify disparities linked to the origin of the data. Fig. 1 visualizes this process, with 'Z' indicating the input noise vector, 'G(Z)' the data generated by the Generator, 'X' the real data, 'G' the Generator, and 'D' the Discriminator.
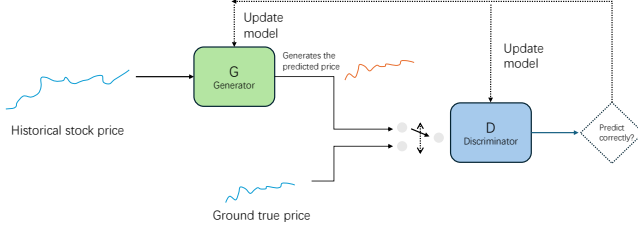

Fig. 1: Structure of Generative Adversarial Network(GAN).

In GAN, both models, denoted as (G, D), are engaged in a continual process of mutual learning, a topic we will delve into in the next section. The GAN model is represented by individual neural networks, with each network operating in opposition to the other. Specifically, x is a sample drawn from the real data distribution $P_{data}(x)$, while z is sampled from a prior distribution $p_z(z)$, such as a uniform or Gaussian distribution. The symbol $E(\cdot)$ represents the expectation operator. $D(x)$ represents the probability of x being sampled from the real data rather than being generated. When the input data is from real data, the discriminator aims to maximize $D(x)$, approaching the value of 1. Conversely, when the input data is generated by the generator $G(z)$, the discriminator endeavors to make $D(G(z))$ approach 0 and generator G tries to make it to 1.

Given the competitive nature of this zero-sum game involving two players, the optimization problem of GAN can be formulated as a minimax problem or the lossing function:

$$\min_G \max_D V(D, G) = \mathbb{E}_{\boldsymbol{x} \sim p_{data}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))] \quad (1)$$

In summary, when training Generative Adversarial Networks (GANs), our objective is to learn the model's parameters through competitive training of the generator and discriminator. The generator's task is to produce realistic data, making it difficult for the discriminator to distinguish, while the discriminator aims to accurately differentiate between real and generated data. During the training process, the generator attempts to deceive the discriminator, while the discriminator strives to enhance its ability to discriminate between real and generated data. Through this competition, the generator and discriminator of GANs gradually improve their performance, ultimately achieving the goal of generating high-quality and realistic data.

B. Transformer

Radford et al. [6] conducted ablation studies to assess the Transformer architecture's effectiveness in comparison to models based on Long Short-Term Memory (LSTM). They observed a significant reduction in performance metrics, such as accuracy and training efficiency, when employing LSTM models. This decline is attributed to LSTMs' inherent limitation in capturing long-distance dependencies due to issues like the vanishing gradient problem, restricting their ability to model complex linguistic details effectively. Additionally, CNN has the same limitation of processing long-term temporal dependencies [31]. In contrast, The adoption of Transformer networks, with their self-attention mechanism, markedly improves the model's capacity to understand extended linguistic structures [32] [33] [34].This advantage stems from the Transformers' ability to weigh the importance of all parts of the input data, regardless of their position, facilitating a deeper comprehension of language and context-based content. The key principle of transformer is attention mechanism.

Classical attention mechanisms, such as traditional, multihead, and temporal attention, play a pivotal role across a broad spectrum of applications. These mechanisms are instrumental in advancing fields like text classification, machine translation, image recognition, and video processing. Each type of attention mechanism offers unique advantages in processing and interpreting complex data structures, thereby contributing significantly to improvements in these areas [7] [8] [35] [36].

In this study, we have employed the aforementioned attention mechanisms to adeptly capture contextual nuances and salient features inherent in historical stock price data.

C. Black-Litterman Framework

The Mean-Variance (MV) model, while foundational in portfolio optimization, has limitations, particularly its sensitivity to input estimates, which can lead to non-robust portfolio allocations. This sensitivity can result in portfolios that are not truly diversified and may overconcentrate in a few assets [37] Moreover, the MV model does not adequately address the asymmetry of stock returns and can underestimate risk during market downturns. These shortcomings justify the need for more robust approaches, such as the Black-Litterman model, which integrates market equilibrium with investor views to enhance portfolio construction [38], aiming for more intuitive and stable asset allocation [39], addressing the MV model's issues by generating portfolios with lower risk, less extreme asset allocations, and higher diversification across asset classes [40].

This equation represents the calculation of the equilibrium return vector, Π, where δ denotes the risk aversion coefficient, Σ is the covariance matrix of asset returns, and w_{marker} are the market capital weights.

$$\Pi = \delta \Sigma w_{market} \quad (2)$$

Ω quantifies the uncertainty associated with the investor's views. It is a diagonal matrix where each diagonal element represents the confidence or variance around each view expressed

by matrix P. The term τ is a scalar that reflects the uncertainty of the equilibrium return vector, Σ is the covariance matrix of returns, and P is a matrix that maps each view onto the assets.

$$\Omega = \text{diag}(\mathsf{P}(\tau\Sigma)P^T) \tag{3}$$

This equation adjusts the original covariance matrix, Σ, to account for the additional uncertainty introduced by the investor's views. The adjusted covariance matrix, $\hat{\Sigma}$, incorporates both the market's inherent volatility and the specific risk perceptions associated with the investor's views.

$$\hat{\Sigma} = \Sigma + \left((\tau\Sigma)^{-1} + \mathsf{P}^{\mathrm{T}}\Omega^{-1}\mathsf{P}\right)^{-1} \tag{4}$$

μ_{BL} denotes the Black-Litterman expected returns, which combine the market equilibrium returns with the investor's views. This is achieved by adjusting the market equilibrium returns, Π, with the investor's specific views on returns, Q, taking into account the confidence in these views as captured by Ω.

$$\mu_{BL} = \left[(\tau\Sigma)^{-1} + P^{'}\Omega^{-1}P\right]\left[(\tau\Sigma)^{-1}\Pi + P^{'}\Omega^{-1}Q\right] \tag{5}$$

$$\max_{w} \frac{(w^T\mu_{BL}) - r_f}{\sqrt{w^T\hat{\Sigma}w}} \tag{6}$$

$$\text{subject to: } \sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0 \tag{7}$$

As is shownd in Formulas (6) and (7), the maximization of the Sharpe ratio is formulated as a convex optimization problem to determine the optimal asset weights in a portfolio. This formulation adapts to the Black-Litterman model's outputs, taking into account the expected returns and the adjusted covariance matrix. In line with common institutional investment practices, our portfolio construction excludes negative weighting to reflect the general avoidance of short selling by institutional investors [5]. Hence, the Sharpe ratio maximization is subject to the constraints that the sum of the weights equals one and that all weights are non-negative. This technique ensures that the portfolio construction is not only aligned with the investor's views but also optimized for risk-adjusted returns.

In evaluating the effectiveness of these methods, it is essential to consider factors such as accuracy in predicting market trends, adaptability to market changes, and the ability to mitigate biases inherent in manual view generation. Studies have shown that portfolios constructed using deep learning-enhanced Black-Litterman models tend to outperform those based on traditional weight allocation schemes. This suggests a significant potential for advanced methods, especially deep learning algorithms, to improve the efficacy of the Black-Litterman model in contemporary portfolio management.

IV. METHODOLOGY

In this section, we delve into the BL-TGAN (Black-Litterman Transformer-GAN) architecture, a sophisticated architecture that marries the predictive prowess of Transformer networks with the generative capabilities of Generative Adversarial Networks (GANs), all within the Black-Litterman framework. The Transformer's role is to analyze time-series data for feature extraction and trend prediction, while the GAN is tailored to refine these predictions, ensuring that the generated investment views are not just a reflection of past patterns but also adaptable to new market dynamics.

The methodology's phases are concisely depicted in Fig. 2 below.

A. Hybrid Combination of Transformer and GAN

Transformer-GAN is the combination of transformer model and GAN architecture. The innovative approach of combining Transformer models and Generative Adversarial Networks (GANs) in the context of financial time series prediction, specifically for predicting stock closing prices, leverages the mechanism of GANs—setting generator and discriminator as Transformer models—rather than generating synthetic data. This novel method enhances prediction accuracy and reliability by integrating the strengths of both architectures in a unique manner.

1) Adaptation of GAN Mechanism: Contrary to conventional uses of GANs, which are typically aimed at creating synthetic data that mimics real datasets, our research repurposes the GAN architecture to directly forecast stock closing prices.

In our novel approach, the GAN generator is designed to generate predictions of stock prices, while the discriminator evaluates the accuracy of these predictions against actual historical data. This paradigm shift leverages the adversarial process of GANs to fine-tune the generator's ability to forecast prices, with the discriminator acting as a dynamic benchmark that guides the generator towards more precise estimations. This adversarial training approach allows the generator to learn and improve its predictions over time, resulting in more accurate and realistic future stock price predictions [12] [41].

2) The Superior Features of Transformer Models: The transformer model, a revolutionary architecture in machine learning, operates on principles that enable it to process sequential data efficiently. Central to its functionality are the concepts of attention mechanisms [32] [33] [34] and positional encoding [42] [43], which allow it to handle dependencies and understand the order in sequential data like financial time series.

In the context of transformer models, this is typically realized through the scaled dot-product attention, which can be mathematically expressed as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{8}$$

Here, Q, K, and V represent the query, key, and value matrices, respectively, derived from the input data, and d_k is
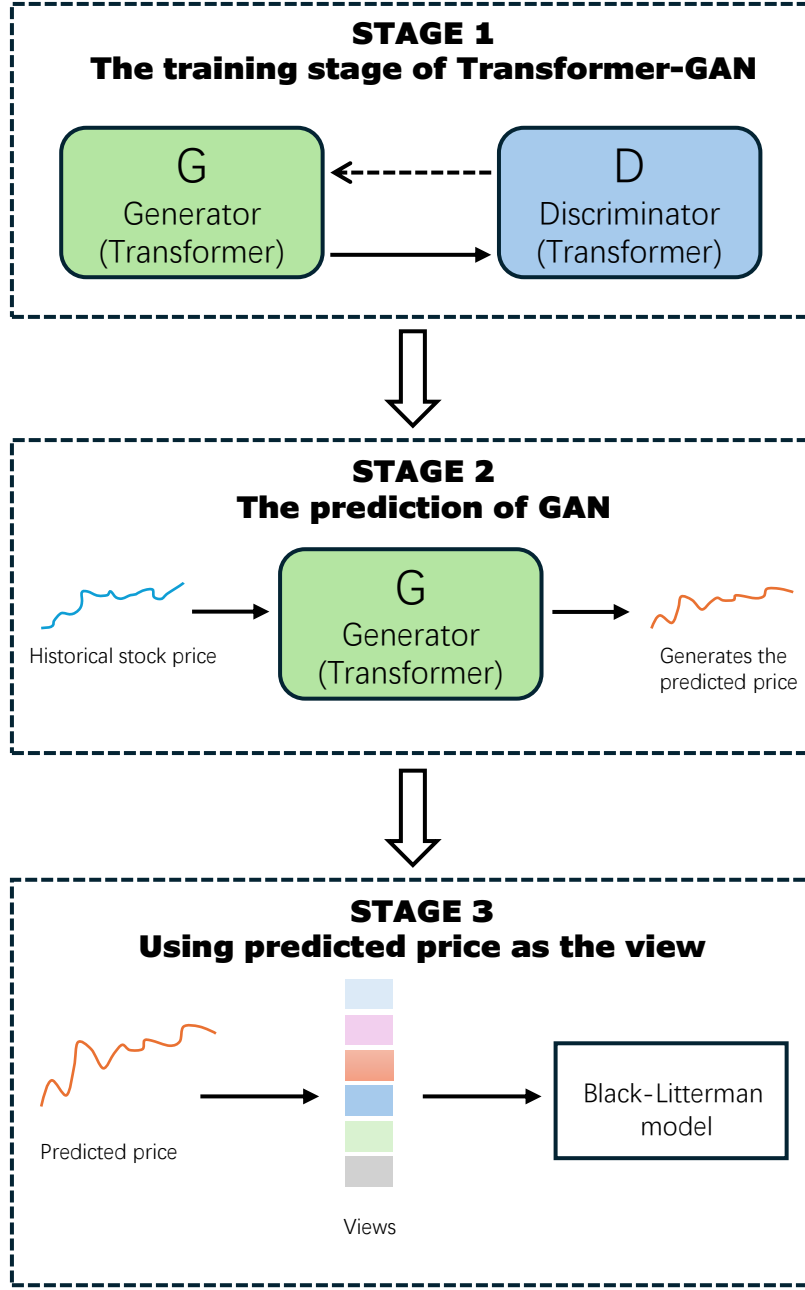
Fig. 2: Architecture of our portfolio construction method

the dimension of the key vector. This formula allows the model to dynamically assign importance to different parts of the input sequence based on their relevance to the current processing context.

Positional encoding is added to the input embeddings to give the model information about the position of the words in the sequence. The positional encoding can be represented as:

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{2i/d_{\text{model}}}}\right) \quad (9)$$

$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{2i/d_{\text{model}}}}\right) \quad (10)$$

In these equations, pos is the position in the sequence, and i is the dimension.

The attention function is then applied in parallel, yielding multiple output values that are concatenated and once again linearly projected. This mechanism is represented as:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, ..., \text{head}_h)W^O \quad (11)$$

where each head is defined as:

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (12)$$

Here, W_i^Q, W_i^K, W_i^V, and W^O are parameter matrices, and h is the number of heads.

The utilization of transformer models in this study is pivotal for generating informed and data-driven views within the Black-Litterman framework. The implementation process is outlined in distinct phases, ensuring a comprehensive and meticulous approach.

3) Transformer-GAN Working Mechanism: In this hybrid model, while the transformer serving as the generator focuses on analyzing historical stock prices to predict future closing prices, the discriminator transformer model evaluates the accuracy of these predictions against real historical data. This competitive setup encourages the predictive model to generate increasingly accurate forecasts, thereby improving its ability to capture the underlying trends and patterns in the data.

The benefits of this hybrid approach include:

- **Enhanced Predictive Accuracy**: The iterative refinement process between the generator and discriminator transformers leads to improved prediction accuracy. The discriminator's challenge compels the generator to produce outputs that are closely aligned with real market behaviors.
- **Robustness to Market Volatility**: The model's exposure to continuous evaluation and challenge enhances its robustness, making it better equipped to handle market volatility and adapt to emerging patterns not explicitly present in the training data.
- **Capturing Latent Market Signals**: GANs are exceptionally adept at distilling latent information within stock price data, information that often eludes traditional analytical models. The generator component of a GAN learns, through its training process, to predict future stock prices by reflecting complex market signals. This enables the effective unveiling of patterns predictive of future market movements.

B. Generating Views with Transformer-GAN architecture

In our approach, the Transformer-GAN architecture processes historical data to forecast future asset returns. The predicted prices p_t are generated from historical prices $p_{t-1}, \ldots, p_{t-5}$. The weekly return will be calculated by the predicted price and structured into a view vector Q, with each element outlining a return prediction for an asset. The views are linked to their respective assets via the construction of matrix P, while the confidence in each view is quantified by the diagonal matrix Ω, typically based on historical volatility.

These elements are integrated into the Black-Litterman model, which are shown in Table I, adjusting market equilibrium returns with the Formula (5). Here, μ_{BL} denotes the adjusted returns, Π the market equilibrium returns, P the matrix mapping views onto assets, Q the vector of views, Σ the covariance matrix of returns, Ω the confidence in the views, and τ the scalar adjusting market covariance. In our method, τ is 0.025. The resulting posterior returns Π are then utilized in a mean-variance optimization to compute the optimal portfolio weights, balancing the Sharpe ratio against risk.

TABLE I: Mathematical Symbols in the Black-Litterman Model

Symbol	Meaning
μ_{BL}	Adjusted returns
Π	Market equilibrium returns
P	Matrix mapping views onto assets
Q	Vector of views
Σ	Covariance matrix of returns
Ω	Confidence in the views
τ	Scalar adjusting market covariance (which is 0.025 here)

C. Assessment Metrics

In evaluating the performance of predictive models, particularly those applied in financial markets like the transformer model for generating views in the Black-Litterman framework, it is crucial to employ accurate and insightful assessment metrics. This study utilizes a suite of metrics to quantify the predictive accuracy and reliability of the model outputs, including Mean Squared Error (MSE), Mean Absolute Error (MAE), and Normalized Mean Squared Error (NMSE). These metrics offer a comprehensive view of model performance, each highlighting different aspects of prediction errors.

- **Mean Squared Error (MSE)**
  The Mean Squared Error (MSE) is a standard metric used to measure the average of the squares of the errors, that is, the average squared difference between the estimated values ($\widehat{y_i}$) and the actual values (y_i). It is defined mathematically as:

$$\text{MSE} = \frac{1}{\zeta} \sum_{i=1}^{\zeta} (y_i - \widehat{y_i})^2 \tag{13}$$

  where ζ represents the total number of observations. MSE is particularly useful for quantifying the magnitude of the prediction error, with lower values indicating better model performance. However, it is sensitive to outliers as it squares the errors before averaging.

- **Mean Absolute Error (MAE)**
  The Mean Absolute Error (MAE) measures the average magnitude of the errors in a set of predictions, without considering their direction. It is calculated as the average of the absolute differences between predicted and actual values:

$$\text{MAE} = \frac{1}{\zeta} \sum_{i=1}^{\zeta} |y_i - \widehat{y_i}| \tag{14}$$

  MAE provides a straightforward interpretation of prediction errors and, unlike MSE, is not unduly influenced by outliers. Lower MAE values signify more accurate predictions.

- **Normalized Mean Squared Error (NMSE)**
  The Normalized Mean Squared Error (NMSE) offers a way to compare the model's error relative to the variance

of the dataset, thereby providing a normalized measure of prediction accuracy:

$$\text{NMSE} = \frac{\sum_{i=1}^{\zeta}(y_i - \widehat{y}_i)^2}{\zeta^* \text{var}(y_i)} \qquad (15)$$

In this equation, $\text{var}(y_i)$ denotes the variance of the actual values. NMSE adjusts the MSE by the variance of the data, making it a relative measure of model performance. Values closer to zero indicate better model accuracy, with negative values suggesting performance worse than the simple mean of the target variable.

V. EXPERIMENT

This section presents a detailed experimental study, our study designs a series of experiments employing real-world financial data. We outline the data sources, the preprocessing steps taken to ensure data quality, and the rigorous evaluation criteria we will apply to assess the model's performance. Through these experiments, we anticipate showcasing our model's ability to enhance portfolio construction and optimization based on black-litterman framework, thus offering valuable contributions to the domain of quantitative finance.

A. Experiment Setup

The experiment comprises several steps, ranging from data simulation to applying the Transformer-GAN model, followed by integrating its output into the Black-Litterman model, and concluding with portfolio optimization.

B. Data Collection and Processing

1) Data Source: As discussed, the dataset provided by Kaggle [44] encompasses historical prices of assets, trading volumes, and various financial metrics. This dataset also includes tweets for the top 25 most-watched stock tickers on Yahoo Finance, spanning from 30-09-2021 to 30-09-2022. This dataset serves as a critical resource for training the transformer models, as well as for other models employed in our controlled experiments. These experiments aim to test the effectiveness of the models in a practical setting.

2) Data Preprocessing: We normalize the stock prices to a range of (0,1) and revert them to their original scale upon obtaining results, which are then utilized to generate views for the Black-Litterman model. The data is partitioned into training and testing sets with proportions of 0.8 and 0.2, respectively. These preprocessing steps are crucial to ensuring the integrity and reliability of our model training and analysis. Subsequently, we conducted a statistical analysis of samples from dataset, the results of several tickers which are presented in Table II.

TABLE II: Statistic Information of Sample from Our Stock Pool(2021-2022)

Stock	Count	Mean	Max	Min	Std
ENPH	251	0.003562	0.246512	-0.123061	0.047608
BX	251	-0.000793	0.084490	-0.100033	0.031066
F	251	-0.000347	0.116674	-0.123242	0.031439
MSFT	251	-0.000491	0.066852	-0.054978	0.019609
GOOG	251	-0.000985	0.077390	-0.058640	0.021730
AAPL	251	0.000228	0.069778	-0.058680	0.020061
VZ	251	-0.001248	0.043521	-0.067352	0.013043
CRM	251	-0.002059	0.098789	-0.117420	0.027698
AMD	251	-0.001185	0.101364	-0.100103	0.037373
KO	251	0.000375	0.038671	-0.069626	0.012234

C. Proposed approach

To demonstrate the comparative effectiveness of our methodology, we utilize Amazon (AMZN) as a case study. The predictive performance of our approach is assessed against other prevalent models, with the findings summarized in Table III.

Within the scope of our model performance analysis, it is evident that the Transformer-GAN model excelled, delivering the most accurate predictions. As delineated in Table III, this model achieved the lowest Mean Absolute Error (MAE) at 0.0373, the smallest Mean Squared Error (MSE) of 0.0025, and the lowest Normalized Mean Squared Error (NMSE) of 0.2088. These results underscore its superior forecasting capability in comparison to the other models evaluated, including standalone CNN, LSTM, Transformer, and their GAN-enhanced counterparts, with the LSTM-GAN model notably underperforming. Furthermore, the CNN-GAN model under-performs relative to both standalone CNN and LSTM models, indicating that not all mechanisms are compatible with or benefit from integration with GAN structure.

TABLE III: Performance Metrics of Predictive Models

Model	MAE	MSE	NMSE
CNN	0.0394	0.0027	0.2237
LSTM	0.0414	0.0032	0.2646
Transformer	0.0385	0.0026	0.2204
CNN-GAN	0.0530	0.0048	0.4003
LSTM-GAN	0.1226	0.0202	1.6920
Transformer-GAN	**0.0373**	**0.0025**	**0.2088**

D. Stock prediction

We use Transformer-GAN to predict the one-day-ahead stock price, and calculate the one-week-ahead return based on the prediction result. Before the prediction, in order to minimize data modification errors, the values were normalized using $MinMaxScaler$. Then, we calculate the predicted return for all stocks, and the results are shown in Table IV. We then select the 10 stocks with the highest predicted returns for inclusion in our portfolio.

TABLE IV: Weekly Return of Stocks

Stock	Predicted Weekly Return
ENPH	0.2157
BX	0.0954
F	0.0786
MSFT	0.0673
GOOG	0.0664
AAPL	0.0420
VZ	0.0401
CRM	0.0351
AMD	0.0295
KO	0.0249
PG	0.0145
NIO	0.0133
NOC	0.0118
TSLA	0.0085
COST	0.0055
ZS	0.0009
AMZN	-0.0040
TSM	-0.0080
DIS	-0.0085
INTC	-0.0103
BA	-0.0133
META	-0.0325
PYPL	-0.0514
XPEV	-0.0627
NFLX	-0.0690

TABLE V: The Market Cap of the 10 Highest Predicted Return

Stock	Market Cap (USD)
ENPH	17,686,683,648
BX	152,683,495,424
F	48,398,815,232
MSFT	3,018,393,387,008
GOOG	1,688,401,346,560
AAPL	2,636,395,315,200
VZ	166,110,707,712
CRM	296,121,597,952
AMD	335,098,675,200
KO	256,677,609,472

The test results depicted in Fig. 3 demonstrate the Transformer-GAN model's proficiency in forecasting stock prices. The model's predictions, shown in green, closely align with the actual stock prices, which are represented in blue. The overlap of the predicted and true values throughout the timeline signifies the model's robust predictive capability. These findings substantiate the model's potential as an effective tool for financial analysts and investors.
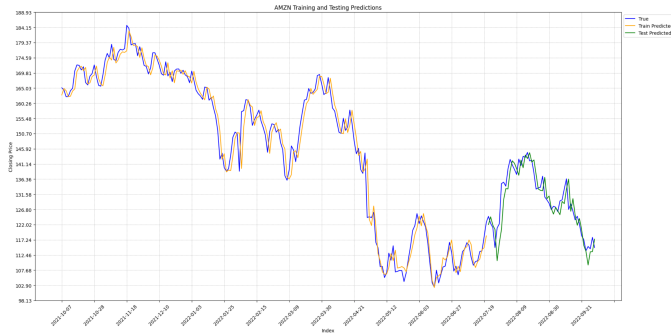
Fig. 3: Closing price predictions by the GAN+TRANS model compared with true values over time.

E. Black-Litterman Model Integration

1) *Implied Equilibrium Return Vector:* Employing market capitalization weights alongside a covariance matrix of returns enabled the calculation of an implied equilibrium return vector, integral to our portfolio optimization strategy. This calculation leverages the proportional market capitals, as detailed in Table V, to inform portfolio optimization. Furthermore, the covariance matrix, denoted by Σ, is illustrated in Fig. 4, providing a visual representation of asset return correlations. Hence, we attain the implied prior returns shown in Fig. 5.
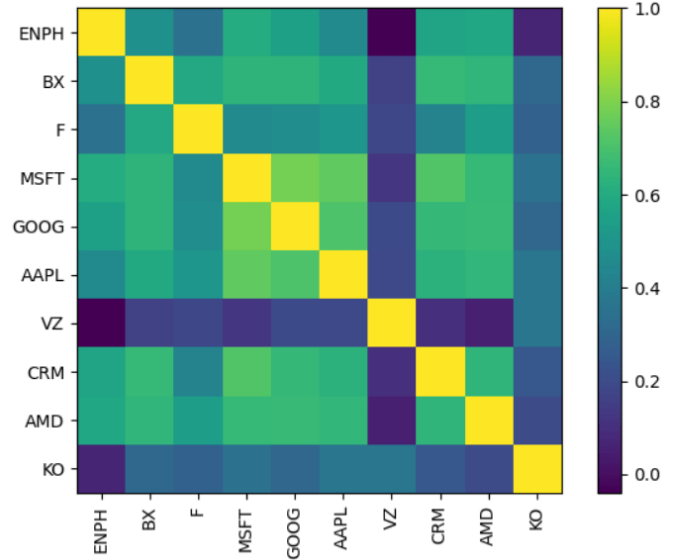

Fig. 4: Covariance Matrix Visualization

2) *Constructing Views* (Q): Applying our Transformer-GAN model to predict the one-week-ahead price, the predicted weekly returns are employed as the views for our analysis, as detailed in Table VI.

F. Portfolio Optimization

In order to validate the efficiency of our proposed method in portfolio construction, we applied various optimization strategies. This includes the traditional mean-variance (MV) approach, the equal-weighted strategy, and the Black-Litterman
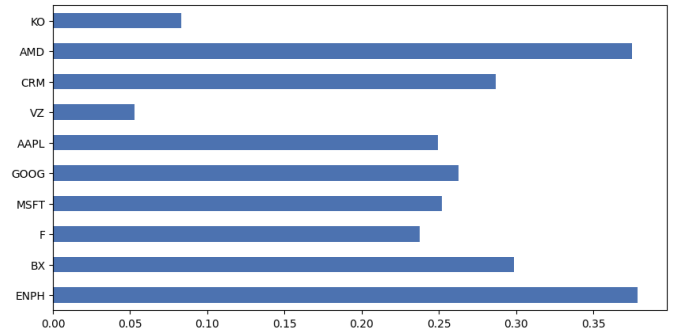

Fig. 5: Implied Prior Returns

TABLE VI: Views for BL model

Stock	View
ENPH	0.2157
BX	0.0954
F	0.0786
MSFT	0.0673
GOOG	0.0664
AAPL	0.0420
VZ	0.0401
CRM	0.0351
AMD	0.0295
KO	0.0249

(BL) model with analyst-derived views. The performance of these strategies is measured against our method to highlight its superior adaptability and effectiveness.

Table VII and Fig. 6 below displays the Prior, Posterior, and Views for a selection of stocks. This data showcases how the BL model adjusts the original market expectations to align more closely with the expectation of our model, allowing for a more informed portfolio construction that potentially leads to enhanced performance.

TABLE VII: Adjusted returns with views

Stock	Prior	Posterior	Views
ENPH	0.378782	0.161047	0.215665
BX	0.298731	0.101838	0.095407
F	0.237640	0.082635	0.078559
MSFT	0.252061	0.101625	0.067298
GOOG	0.262989	0.102187	0.066403
AAPL	0.249424	0.094659	0.041993
VZ	0.053028	0.034654	0.040120
CRM	0.286891	0.090919	0.035144
AMD	0.374907	0.111667	0.029542
KO	0.083053	0.035655	0.024880
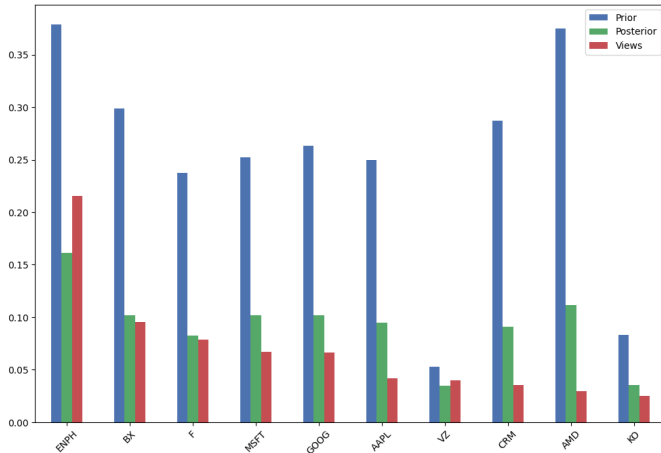
Fig. 6: Adjusted returns with views

Fig. 7 illustrates the weights derived from our method, while Table VIII displays the weights calculated using our method, mean-variance and equal-weighted strategies, respectively.

Fig. 8, 9, and 10 display the returns for various holding periods under different optimization strategies. Fig. 8 illustrates the returns using our method, consistently outperforming
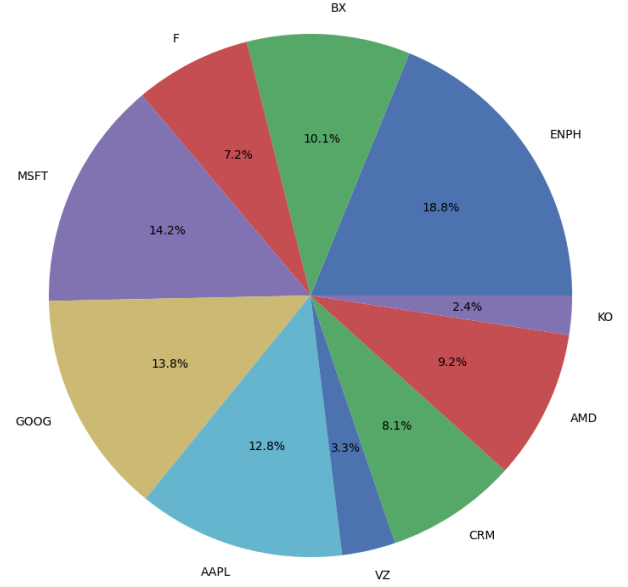

Fig. 7: Weights (Our method)

TABLE VIII: Portfolio Weights by Different Optimization Models

Stock	Our Method	Mean-Variance	Equal Weights
AAPL	12.8%	12.0%	10.0%
MSFT	14.2%	11.8%	10.0%
GOOG	13.8%	12.1%	10.0%
ENPH	18.8%	12.4%	10.0%
KO	2.4%	3.5%	10.0%
AMD	9.2%	14.5%	10.0%
CRM	8.1%	11.7%	10.0%
VZ	3.3%	2.0%	10.0%
BX	10.1%	11.4%	10.0%
F	7.2%	8.4%	10.0%

the benchmarks, particularly for longer holding periods. This is in contrast to Fig. 9 and 10, which represent the MV and equal-weighted strategies, respectively. For the Black-Litterman strategy, Fig. 11 and 12 present outcomes generated based on views from analysts. These figures show a competitive start but generally underperform compared to our method, especially as the holding period increases. In the later stages, the observed minimal discrepancy between the market return and the return generated by our methodology can be attributed to typical market fluctuation which is a common situation.
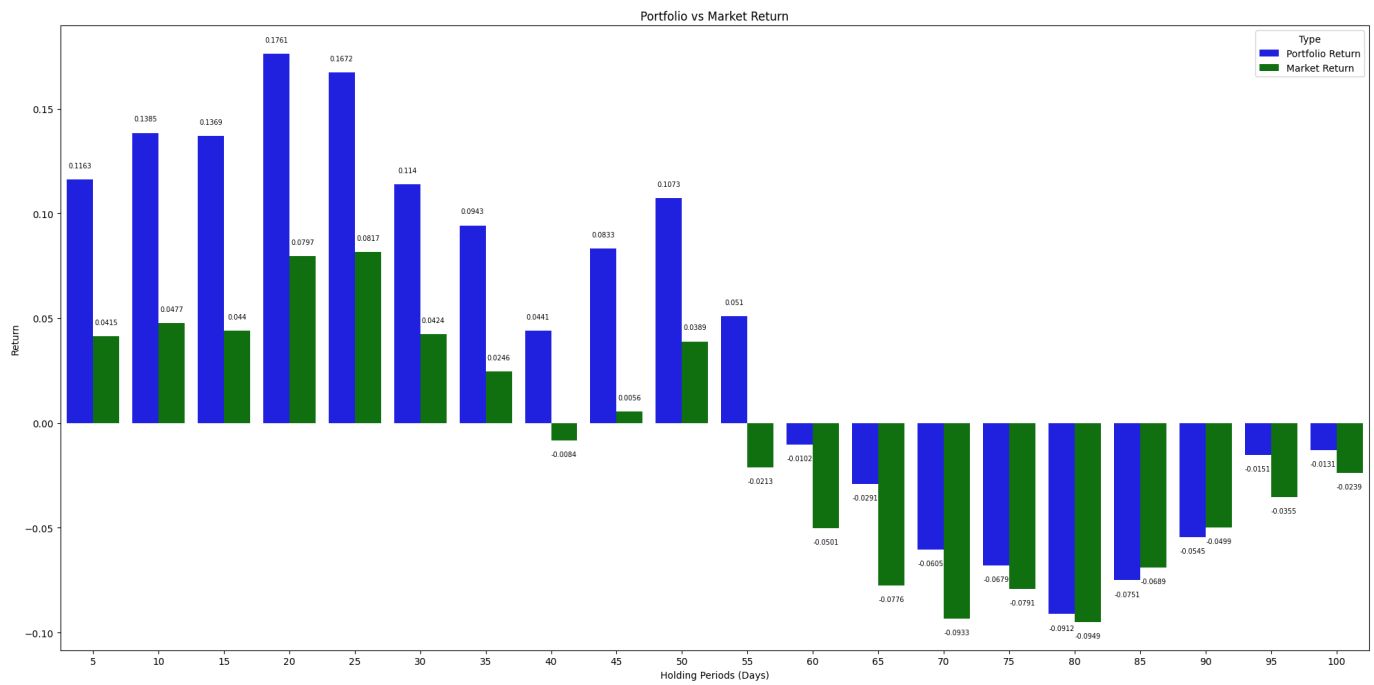
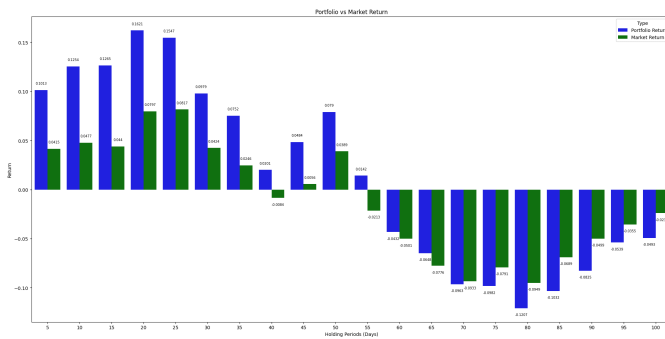Fig. 8: Returns comparison for different periods (Our method)

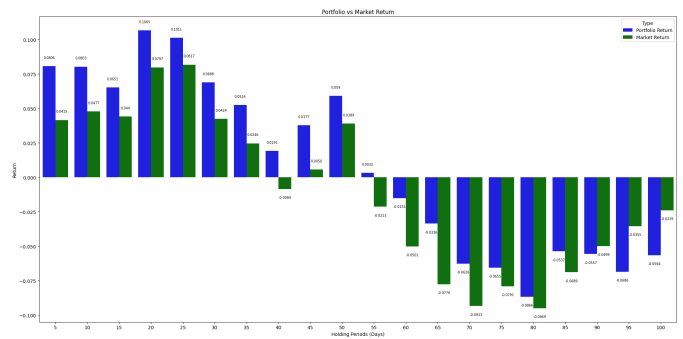Fig. 9: Returns comparison for different periods (MV)

Fig. 11: Returns comparison for different periods (BL:view from analyst 1)
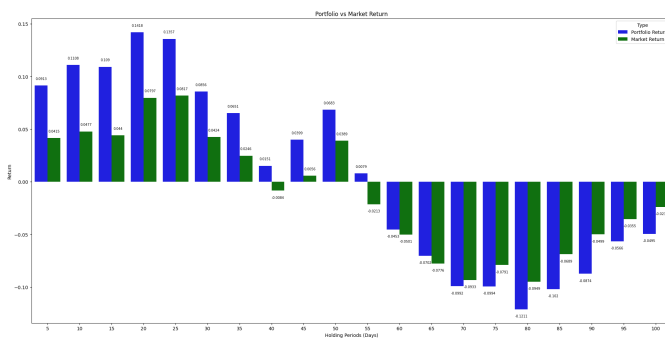

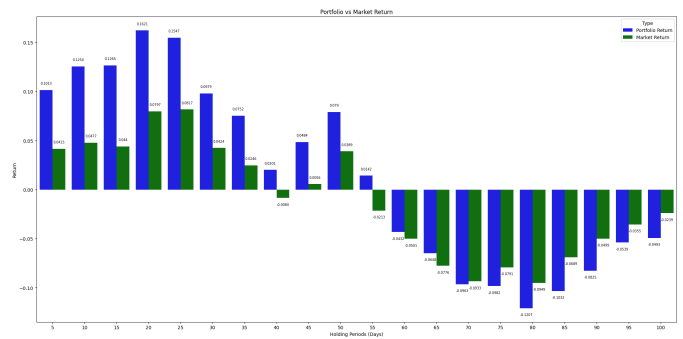Fig. 10: Returns comparison for different periods (Equal weights)

Fig. 12: Returns comparison for different periods (BL:view from analyst 2)

The graphical representation in Fig. 13 provides a comprehensive overview of portfolio returns over various time

periods, contrasting the performance of different investment strategies against actual market returns. Our method, indicated by the **Blue** line which is marked as **'BL'**, exhibits robust returns that surpass the performance of all strategies for the majority of the observed periods. Notably, in the later stages, the market return, illustrated by the orange line, outperforms all the strategies, including ours. This observation is consistent with the typical market behavior, where long-term fluctuations may favor the actual market returns over modeled strategies [45] [46] [47]. Such trends underscore the challenges in outperforming the market consistently, especially in later periods where market dynamics may change rapidly. Despite this, the superior performance of our model in the earlier phases demonstrates its potential to provide investors with significant insights for strategic decision-making in asset management.

Fig. 14 offers a detailed visualization of portfolio returns corresponding to four distinct scenarios, each scenario characterized by a set of randomly selected entry points. The ensemble of these subfigures demonstrates the robust performance of the proposed investment strategy across a spectrum of stochastic market conditions. Despite the inherent unpredictability of entry points, the strategy exhibits a persistent outperformance, highlighting its potential for strategic asset management.

VI. Conclusion

Recent advancements in AI, particularly the popularity of attention mechanisms and the use of LSTM and CNN models for financial forecasting, have significantly improved the predictive accuracy of stock price movements. Building upon these advancements, our study introduces a novel approach by integrating Transformer models—known for their exceptional ability to capture long-range dependencies within data—with Generative Adversarial Networks (GANs). This combination allows our model to not only predict stock prices with high accuracy but also validate these predictions against actual historical data.

The implications of our findings extend beyond the academic sphere, offering tangible benefits to portfolio managers and financial analysts. By incorporating our Transformer-GAN model, practitioners can achieve more accurate predictions of market movements and asset prices, facilitating better decision-making and risk management. Theoretically, this research enriches the dialogue on the utility of machine learning in finance, particularly highlighting the benefits of combining different types of neural networks to address specific financial tasks.

In conclusion, our transformative approach leverages a Transformer-GAN model to generate predictive views for the Black-Litterman model, achieving novel results in portfolio optimization. This integration not only enhances predictive accuracy but also innovates on traditional financial models by incorporating cutting-edge AI technologies. Our findings represent a significant step forward in the application of machine learning in finance, offering a sophisticated method for asset allocation that aligns with modern investment strategies and market expectations.

Despite its promising outcomes, our approach faces challenges, including sensitivity to market volatility, data overfitting risks, scalability concerns due to large data dependencies, and GANs' training instability leading to convergence issues. Innovative solutions are needed to enhance predictive accuracy while maintaining efficiency. Future work may explore hybrid deep learning models to address these limitations, extend the model's application across various financial instruments and markets, and improve computational efficiency for broader industry adoption.

References

[1] F. Black and R. Litterman, "Asset allocation: combining investor views with market equilibrium," *Goldman Sachs Fixed Income Research*, vol. 115, no. 1, pp. 7–18, 1990.

[2] ——, "Global portfolio optimization," *Financial Analysts Journal*, vol. 48, no. 5, pp. 28–43, 1992. [Online]. Available: https://doi.org/10.2469/faj.v48.n5.28

[3] S. Selvin, R. Vinayakumar, E. Gopalakrishnan, V. K. Menon, and K. Soman, "Stock price prediction using lstm, rnn and cnn-sliding window model," in *2017 international conference on advances in computing, communications and informatics (icacci)*. IEEE, 2017, pp. 1643–1647.

[4] H. Rezaei, H. Faaljou, and G. Mansourfar, "Intelligent asset allocation using predictions of deep frequency decomposition," *Expert Systems with Applications*, vol. 186, p. 115715, 2021.

[5] R. Barua and A. K. Sharma, "Dynamic black litterman portfolios with views derived via cnn-bilstm predictions," *Finance Research Letters*, vol. 49, p. 103111, 2022.

[6] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever *et al.*, "Improving language understanding by generative pre-training," 2018.

[7] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[8] R. Fan, W. Chu, P. Chang, J. Xiao, and A. Alwan, "An improved single step non-autoregressive transformer for automatic speech recognition," *arXiv preprint arXiv:2106.09885*, 2021.

[9] Q. Zhang, C. Qin, Y. Zhang, F. Bao, C. Zhang, and P. Liu, "Transformer-based attention network for stock movement prediction," *Expert Systems with Applications*, vol. 202, p. 117239, 2022.

[10] C. Wang, Y. Chen, S. Zhang, and Q. Zhang, "Stock market index prediction using deep transformer model," *Expert Systems with Applications*, vol. 208, p. 118128, 2022.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, Eds., vol. 27. Curran Associates, Inc., 2014. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf

[12] M. Diqi, M. E. Hiswati, and A. S. Nur, "Stockgan: robust stock price prediction using gan algorithm," *International Journal of Information Technology*, vol. 14, no. 5, pp. 2309–2315, 2022.

[13] S. R. Polamuri, K. Srinivas, and A. K. Mohan, "Multi-model generative adversarial network hybrid prediction algorithm (mmgan-hpa) for stock market prices prediction," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 9, pp. 7433–7444, 2022.

[14] X. Yin, Y. Han, H. Sun, Z. Xu, H. Yu, and X. Duan, "Multi-attention generative adversarial network for multivariate time series prediction," *IEEE Access*, vol. 9, pp. 57351–57363, 2021.
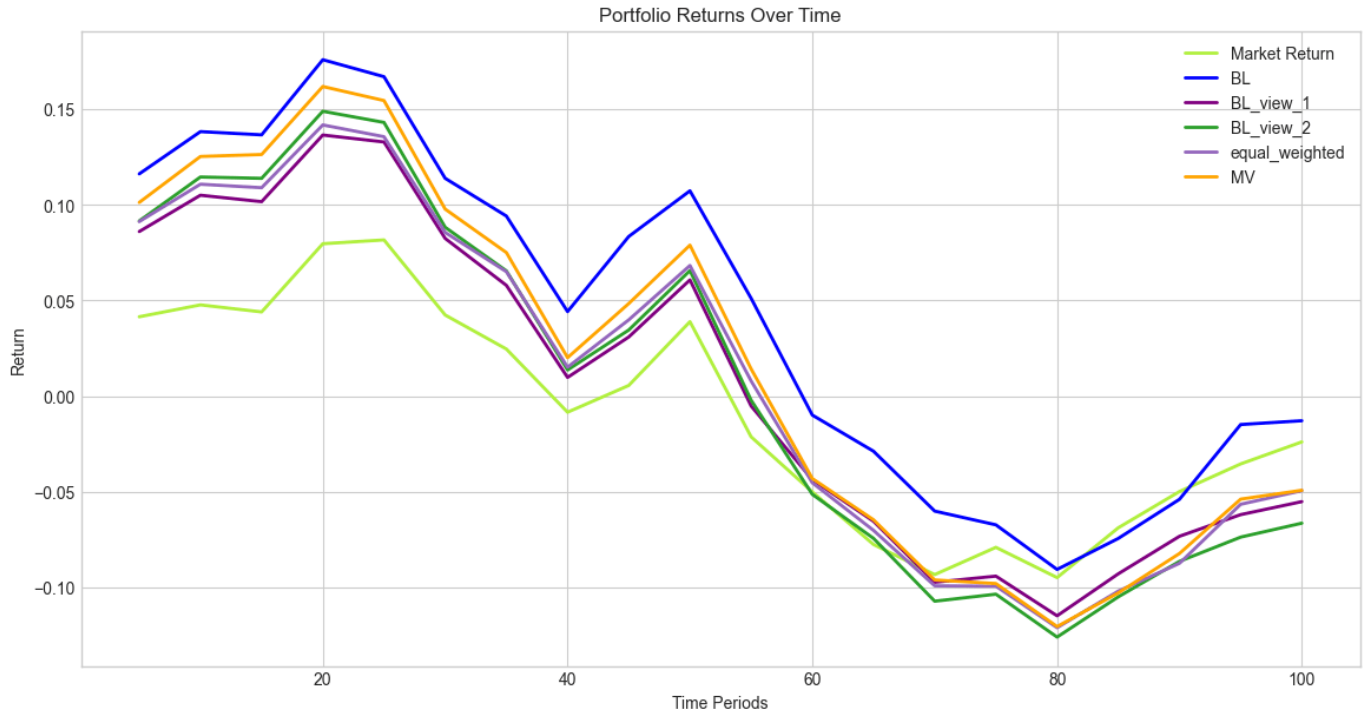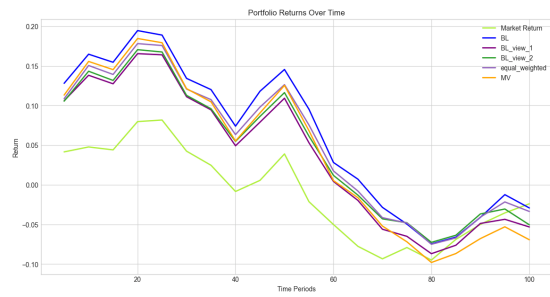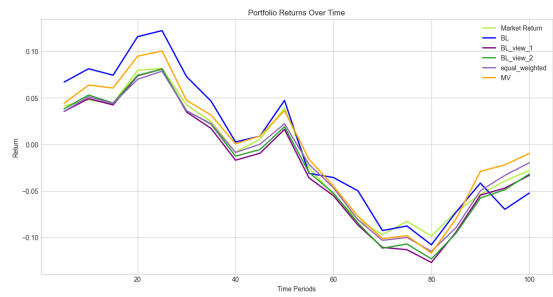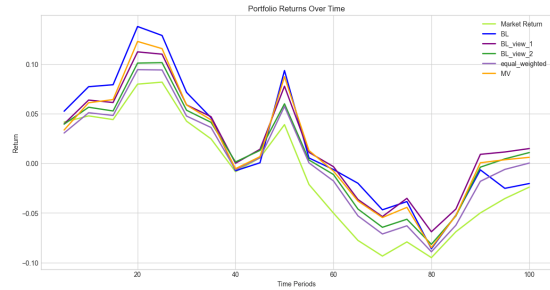
Fig. 13: Comparative Analysis of Portfolio Returns Across Different Investment Strategies
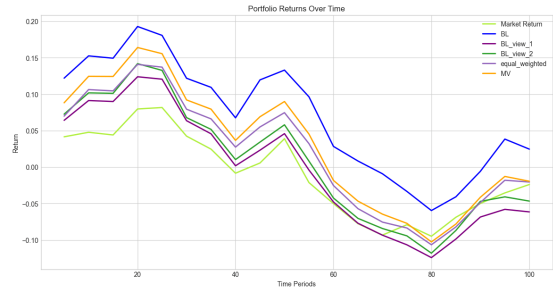

(a) Random Entry Analysis I

(b) Random Entry Analysis II

(c) Random Entry Analysis III

(d) Random Entry Analysis IV

Fig. 14: Stochastic Entry Points and Portfolio Performance

[15] A. Ferchichi, A. B. Abbes, V. Barra, M. Rhif, and I. R. Farah, "Multi-attention generative adversarial network for multi-step vegetation indices forecasting using multivariate time series," *Engineering Applications of Artificial Intelligence*, vol. 128, p. 107563, 2024.

[16] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Applied soft computing*, vol. 90, p. 106181, 2020.

[17] J. M.-T. Wu, Z. Li, G. Srivastava, J. Frnda, V. G. Diaz, and J. C.-W. Lin, "A cnn-based stock price trend prediction with futures and historical price," in *2020 International Conference on Pervasive Artificial Intelligence (ICPAI)*. IEEE, 2020, pp. 134–139.

[18] A. Vidal and W. Kristjanpoller, "Gold volatility prediction using a cnn-lstm approach," *Expert Systems with Applications*, vol. 157, p. 113481, 2020.

[19] L. Di Persio, O. Honchar *et al.*, "Recurrent neural networks approach to the financial forecast of google assets," *International journal of Mathematics and Computers in simulation*, vol. 11, pp. 7–13, 2017.

[20] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A cnn-lstm-based model to forecast stock prices," *Complexity*, vol. 2020, pp. 1–10, 2020.

[21] W. Lu, J. Li, J. Wang, and L. Qin, "A cnn-bilstm-am method for stock price prediction," *Neural Computing and Applications*, vol. 33, no. 10, pp. 4741–4753, 2021.

[22] K. Wu, J. Wu, L. Feng, B. Yang, R. Liang, S. Yang, and R. Zhao, "An attention-based cnn-lstm-bilstm model for short-term electric load forecasting in integrated energy system," *International Transactions on Electrical Energy Systems*, vol. 31, no. 1, p. e12637, 2021.

[23] L. Sun, W. Xu, and J. Liu, "Two-channel attention mechanism fusion model of stock price prediction based on cnn-lstm," *Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 5, pp. 1–12, 2021.

[24] J. Zhang, L. Ye, and Y. Lai, "Stock price prediction using cnn-bilstm-attention model," *Mathematics*, vol. 11, no. 9, p. 1985, 2023.

[25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[26] C. Wang, Y. Chen, S. Zhang, and Q. Zhang, "Stock market index prediction using deep transformer model," *Expert Syst. Appl.*, vol. 208, p. 118128, 2022. [Online]. Available: https://doi.org/10.1016/j.eswa.2022.118128

[27] Q. Zhang, C. Qin, Y. Zhang, F. Bao, C. Zhang, and P. Liu, "Transformer-based attention network for stock movement prediction," *Expert Syst. Appl.*, vol. 202, p. 117239, 2022. [Online]. Available: https://doi.org/10.1016/j.eswa.2022.117239

[28] Y. Hong, U. Hwang, J. Yoo, and S. Yoon, "How generative adversarial networks and their variants work: An overview," *ACM Computing Surveys (CSUR)*, vol. 52, no. 1, pp. 1–43, 2019.

[29] X. Zhang, F. Li, and N. U. Shah, "Hidden knowledge recovery from gan-generated single-cell rna-seq data," *bioRxiv*, pp. 2023–11, 2023.

[30] B. J. B. Rani *et al.*, "Survey on applying gan for anomaly detection," in *2020 International Conference on Computer Communication and Informatics (ICCCI)*. IEEE, 2020, pp. 1–5.

[31] Z. Zeng, R. Kaur, S. Siddagangappa, S. Rahimi, T. Balch, and M. Veloso, "Financial time series forecasting using cnn and transformer," *arXiv preprint arXiv:2304.04912*, 2023.

[32] C. D. Manning, K. Clark, J. Hewitt, U. Khandelwal, and O. Levy, "Emergent linguistic structure in artificial neural networks trained by self-supervision," *Proceedings of the National Academy of Sciences*, vol. 117, no. 48, pp. 30 046–30 054, 2020.

[33] K. Ameri, M. Hempel, H. Sharif, J. Lopez, and K. Perumalla, "Impact of grammar on language model comprehension," in *2023 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2023, pp. 19–24.

[34] J. Ferrando, G. I. Gállego, I. Tsiamas, and M. R. Costa-jussà, "Explaining how transformers use context to build predictions," *arXiv preprint arXiv:2305.12535*, 2023.

[35] V. Gabeur, C. Sun, K. Alahari, and C. Schmid, "Multi-modal transformer for video retrieval," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IV 16*. Springer, 2020, pp. 214–229.

[36] C. Yan, Y. Tu, X. Wang, Y. Zhang, X. Hao, Y. Zhang, and Q. Dai, "Stat: Spatial-temporal attention mechanism for video captioning," *IEEE transactions on multimedia*, vol. 22, no. 1, pp. 229–241, 2019.

[37] W. Bessler, H. Opfer, and D. Wolff, "Multi-asset portfolio optimization and out-of-sample performance: an evaluation of black–litterman, mean-variance, and naïve diversification approaches," *The European Journal of Finance*, vol. 23, no. 1, pp. 1–30, 2017.

[38] T. Teplova, M. Evgeniia, Q. Munir, and N. Pivnitskaya, "Black-litterman model with copula-based views in mean-cvar portfolio optimization framework with weight constraints," *Economic Change and Restructuring*, vol. 56, no. 1, pp. 515–535, 2023.

[39] M. Adelmann, "An improvement of the global minimum variance portfolio using a black-litterman approach," *Available at SSRN 2779169*, 2016.

[40] F. J. Fabozzi, S. M. Focardi, and P. N. Kolm, "Incorporating trading strategies in the black-litterman framework," *The Journal of Trading*, vol. 1, no. 2, pp. 28–37, 2006.

[41] D. Ma, D. Yuan, M. Huang, and L. Dong, "Vgc-gan: a multi-graph convolution adversarial network for stock price prediction," *Expert Systems with Applications*, vol. 236, p. 121204, 2024.

[42] V. Shiv and C. Quirk, "Novel positional encodings to enable tree-based transformers," *Advances in neural information processing systems*, vol. 32, 2019.

[43] A. Kazemnejad, I. Padhi, K. Natesan Ramamurthy, P. Das, and S. Reddy, "The impact of positional encoding on length generalization in transformers," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[44] Equinxx, "Stock Tweets for Sentiment Analysis and Prediction," https://www.kaggle.com/datasets/equinxx/stock-tweets-for-sentiment-analysis-and-prediction/data, 2023, accessed: yyyy-mm-dd.

[45] H. Svedsater, N. Karlsson, and T. Garling, "Momentum trading, disposition effects and prediction of future share prices: an experimental study of multiple reference points in responses to short-and long-run return trends," *Applied Financial Economics*, vol. 19, no. 8, pp. 595–610, 2009.

[46] M. Malin and G. Bornholt, "Long-term return reversal: Evidence from international market indices," *Journal of International Financial Markets, Institutions and Money*, vol. 25, pp. 1–17, 2013.

[47] J. Y. Campbell and L. M. Viceira, "The term structure of the risk–return trade-off," *Financial Analysts Journal*, vol. 61, no. 1, pp. 34–44, 2005.