TIC Quantitative Trading

FROM QUALITATIVE VIEWS TO QUANTITATIVE PORTFOLIOS: A LANGUAGE-MODEL-AUGMENTED BLACK–LITTERMAN

By Catarina Oliveira

By Matilde Silva

Portfolio construction bridges mathematical precision and human intuition, as investors rely on quantitative frameworks while simultaneously incorporating subjective beliefs about future market behavior. Most conventional portfolio models, however, provide limited scope for integrating such investor views in a flexible manner, focusing mostly on exact calculations instead of qualitative information. Moreover, building an effective portfolio is inherently challenging due to the sheer number of investment opportunities and the fundamental difficulty of forecasting future returns. Investors must navigate incomplete information, balancing risk, expected return and subjective insights in markets that are constantly evolving.

In response to these challenges, the Black-Litterman framework arises as a solution that balances inventor views with the objective returns from the market in equilibrium. We propose a flexible implementation that leverages large language models (LLMs) to translate user input into a structured set of views compatible with the Black–Litterman framework, ultimately generating an optimized portfolio.

Classic Black–Litterman Implementation

The Mean–Variance Paradigm and Its Limitations

The foundations of modern portfolio theory were laid by Harry Markowitz, through his mean-variance model. In theory, investors could identify an efficient frontier and select portfolios that maximize expected return for a given level of risk, providing a quantitative solution to the asset allocation problem. where ρ_{ij} is the Pearson correlation between the log-returns of the two assets.

Despite its elegance, the mean-variance framework suffers from several shortcomings; estimates of returns and volatilities are derived from historical data, and the resulting portfolio weights are highly sensitive to estimation errors, which produces economically unintuitive and unstable portfolios. Additionally, this model provides no explicit mechanism for incorporating investor views.

These limitations motivated the search for a more robust and intuitive framework for portfolio construction.

The Black-Litterman Model

In response to the shortcomings of the mean-variance optimization, Fischer Black and Robert Litterman developed the Black–Litterman (BL) model. Their approach represents a significant conceptual shift, by allowing investors to incorporate their subjective views in an objective and quantitative manner. At its core, the BL model blends three major components: **equilibrium expected returns**, inferred from a reference market portfolio; **estimation uncertainty** and **investor views**.

The result is a set of posterior expected returns that balance market equilibrium with investor conviction, producing portfolios that are more stable, intuitive, and aligned with economic reasoning.

Equilibrium Returns

At the BL model foundation lies a general market equilibrium model (or a benchmark portfolio as a proxy), which assumes that investors are risk-averse and seek to maximize a utility function which takes the following expression:

$$U(x) = x'\Pi - \frac{\delta}{2}x'\Sigma x$$

$$= x'(\mu - r_f) - \frac{\delta}{2}x'\Sigma x$$

Here, χ denotes the vector of portfolio weights; δ is the coefficient of risk aversion; Σ is the covariance matrix of asset returns, which captures how assets relate and move together; and Π represents the vector of expected returns above the risk-free rate, given a market at equilibrium, where μ is the expected returns vector and r_f the risk-free rate.

TIC Quantitative Trading

This expression reflects the difference between reward and risk, symbolizing that a portfolio is attractive if it offers high expected returns, but its attractiveness is reduced if it is very risky. Consequently, the utility functionis maximized to identify allocations that deliver the greatest return per unit of risk, consistent with the assumption that investors are risk averse.

Uncertainty in the Reference Model

The BL model takes into account how uncertain are the returns of each asset, while taking into account the likelihood of other assets' estimates being wrong. This error in the equilibrium expected returns is specified by the covariance matrix of the estimation errors, Σ_ϵ which describes how noise the equilibrium returns are with the following expression:

$$\Sigma_\epsilon = \tau\Sigma$$

The scalar parameter τ value is model-dependent, however, it is commonly defined as $1/T$ with T denoting the length of the estimation window.

Modeling Investor Views

The most distinctive feature of the BL framework is its treatment of investor views. These views are expressed through P, a matrix that identifies the assets involved in each view, and the vector Q, specifies the corresponding expected returns. Additionally, the uncertainty associated with investor views is captured by the covariance matrix, Ω, allowing for varying degrees of confidence across different views. Each row of P represents a portfolio; absolute views are encoded by assigning a value of 1 to the relevant asset, while relative views are constructed by assigning a value of 1 to one asset and -1 to another. The corresponding element of Q captures the expected return implied by the view.

The Black-Litterman Expected Returns

By combining equilibrium returns, estimation uncertainty and investor views, the Black-Litterman model yields the posterior vector of expected excess returns:

$$\Pi_{BL} = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q]$$

Black-Litterman implementation

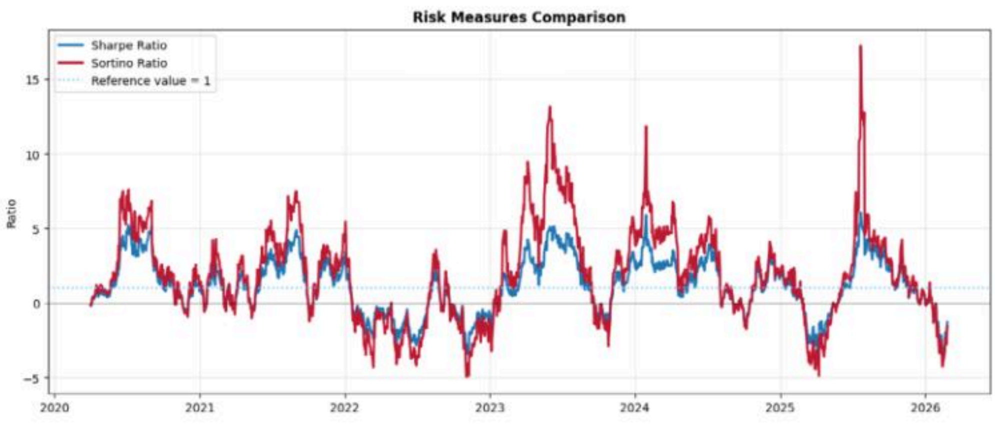
Consider an investor seeking exposure to the largest and most influential companies in the S&P500, selecting Apple (AAPL), Microsoft (MSFT), Alphabet (GOOG/GOOGL), Amazon (AMZN), and NVIDIA (NVDA). Together, these firms dominate market capitalization and play a central role in shaping broader index performance, while capturing key shifts in technology, innovation, and consumer demand. The investor expects Alphabet to deliver an absolute growth of 10% in relation to the market in equilibrium, Microsoft to outperform Apple by 2 percentage points, and Amazon to underperfom NVIDIA by 3 percentage points.

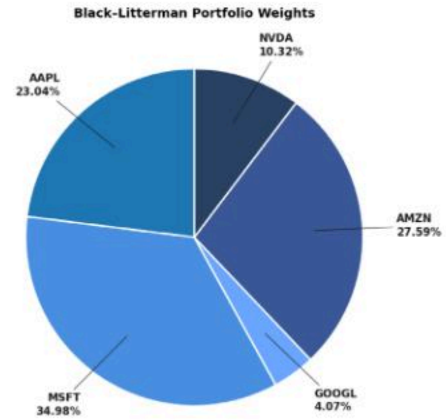
These beliefs can be incorporated directly within the Black–Litterman framework. Keeping the assets ordered as listed above, the investor's views are encoded mathematically in the following matrices:

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

$$Q = \begin{bmatrix} 0.10 \\ 0.02 \\ -0.03 \end{bmatrix}$$

Through the computations described previously, these views are combined with market equilibrium returns, producing the desired portfolio and its expected returns.





This graph portraits the Sharpe and Sortino ratios for a rolling window of 3 months, representing the evolution of the risk-adjusted performance of the created portfolio.

Black Litterman LLM Augmentation

The Black–Litterman model requires numerical inputs, yet investor views are initially expressed in natural language and must therefore be translated into their quantitative form before they can be incorporated into the framework. The translation of informal language into objective data is prone to inconsistency, and difficult to scale across users or scenarios and therefore constitutes a key obstacle in the practical application of the Black-Litterman framework. The introduction of Large Language Models (LLMs) addresses this problem by providing a more reliable mechanism to convert natural language prompts into structured quantitative inputs, enhancing practical usability and making the framework more accessible to a broader range of users, regardless of their technical or quantitative background.

System Architecture

The proposed pipeline consists of three layers: a user interaction layer, a GPT-based interpretation layer, and a portfolio optimization layer. Users express views in natural language along with confidence levels. The language model structures these inputs into formal representations that are directly used by the Black–Litterman model.

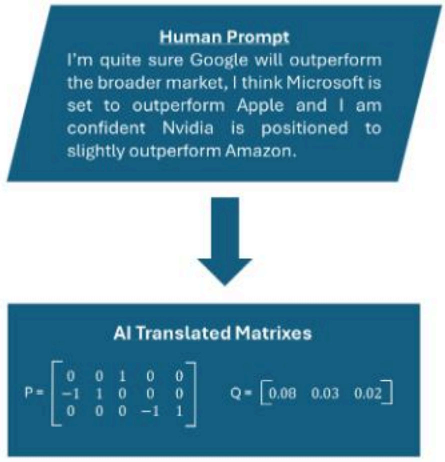
Role of the Language Model

In the proposed pipeline, the language model does not directly participate in portfolio optimization. Instead, its function is strictly confined to the interpretation, validation and finally, structuring of investor views, prior to their incorporation into the Black–Litterman framework. The LLM acts as an interface layer between unstructured human input and the formal mathematical representation required by the model.

Investor views are typically expressed in natural language and often exhibit ambiguity or inconsistency. The language model is used to translate these qualitative statements into structured representations, such as relative or absolute return views and confidence levels and to evaluate the consistency and validity of these views. This translation process enforces consistency, ensuring that views are non-contradictory and compatible with the constraints of the Black–Litterman model.

Practical Application (Using ChatGPT - 4o)

When comparing the P and Q matrices we obtained in the first part of the article and just now, we can see how through a simple natural-language prompt, a clearly instructed LLM can generate P and Q matrices that are very similar to those that were created manually (which is the equivalent of a fully objective and technical prompt).



THE FUTURE

At this stage, we have successfully developed the foundational pipeline that integrates natural-language investor views with the Black–Litterman framework through LLM augmentation. Currently, our focus is on refining the translation of qualitative inputs into structured P and Q matrices, validating their consistency, and testing the resulting portfolios on selected equity datasets. This work positions us to deliver a robust, flexible tool for portfolio construction that bridges human intuition and quantitative rigor, laying the groundwork for further experimentation and practical deployment.

Article written in March and updated on April.