UNIVERSITY
OF OSLO

Bachelor's thesis

Forecast-Informed Portfolio Optimization: A Machine Learning Approach to Strategic Asset Allocation

Storm Hovland

STK-MAT2011 – Project work in finance, insurance, risk and data analysis
10 ECTS study points

Supervised by Gudmund Horn Hermansen

Department of Mathematics
Faculty of Mathematics and Natural Sciences

Spring 2025

Contents

Abstract

This thesis investigates the integration of machine learning and modern portfolio theory to enhance equity investment strategies. Using an XGBoost soft classifier, the model forecasts the probability that individual stocks will exceed a 2% return over a 20-day horizon. A rich set of technical indicators is engineered and optimized to improve prediction performance. The resulting probability estimates are used to identify the most promising assets across sectors, based on classification metrics. These assets are then used for portfolio construction. Both the traditional Markowitz mean-variance framework and the Black-Litterman model are implemented, with the latter incorporating market capitalization data and model-based investor views. The best performing portfolios, selected based on Sharpe ratio, return, and volatility, are backtested against benchmarks. The results suggest that combining machine learning forecasts with portfolio theory can improve asset allocation and deliver more robust investment performance.

1 Introduction

1.1 Background and Motivation

Portfolio optimization has long been recognized as a fundamental problem in finance. The objective is to construct a collection of assets that balance risk and return in the best possible way, typically by allocating capital among a set of financial assets. Since the pioneering work of Markowitz (1952) and the development of the mean-variance model, portfolio theory has continuously evolved to incorporate more sophisticated methods of asset selection and allocation. In recent years, the growing availability of financial data and advances in machine learning have opened opportunities for improving investment strategies and analysis. Machine learning models, particularly ensemble methods such as XGBoost, have demonstrated strong capabilities of capturing complex, nonlinear patterns in financial markets, proposed by Zhang (2022). Despite the promise of machine learning, the field of stock price prediction contains some limitations. Financial markets are inherently noisy and influenced by numerous unpredictable factors. Therefore, accurate stock return predictions are difficult, and according to Kazeem (2023), the integration of such predictions into portfolio construction remains an open research challenge.

1.2 Problem Statement and Objective

While machine learning models such as XGBoost can improve stock market forecasting, several challenges persist:

- **Market volatility:** Financial markets are highly volatile, meaning asset prices can fluctuate significantly due to new information, investor sentiment, or macroeconomic shocks. This volatility introduces noise into the data, making it difficult for machine learning models to distinguish between short-term randomness and meaningful patterns. As a result, even advanced algorithms like XGBoost may struggle to produce consistent and accurate predictions, especially during periods of market stress or regime shifts.

- **Integration with portfolio theory:** Even when probabilistic forecasts are available, integrating them systematically into portfolio optimization, while maintaining diversification and controlling risk is a nontrivial task.

The objective of this thesis is to develop a systematic methodology that addresses these challenges. By combining machine learning-based stock selection with portfolio optimization frameworks like the Markowitz and Black-Litterman models, the aim is to construct investment portfolios that outperform standard benchmarks in terms of risk-adjusted returns. An overview of the full workflow from prediction to optimization and evaluation, is visualized in Figure 2, providing a visual guide to the methodological structure.

2 Research Methodology

2.1 XGBoost Soft Classification

Machine Learning has become increasingly popular in finance due to its ability to capture complex, non-linear relationships in financial time series, as demonstrated by Zhang (2022). In this project, the XGBoost (eXtreme Gradient Boosting) (GeeksforGeeks 2023) algorithm is employed as a **soft classifier** to estimate the probability that a stock will yield a return exceeding 2% over a 20-day horizon. The model outputs a value between 0 and 1, that can be interpreted as the model‘s confidence in the positive class (i.e., the stock outperforming the 2% threshold). These soft scores enable flexible decision-making and ranking of assets by expected performance, which is especially important when constructing portfolios based on predicted likelihoods rather than binary signals. XGBoost builds an ensemble of decision trees in a **stage-wise additive** manner (GeeksforGeeks 2023), where each new tree corrects the residuals of the previous ones. This allows it to efficiently model complex interactions and achieve high predictive performance. Let $x_i \in \mathbb{R}^m$ denote the input feature vector for observation i, where each component corresponds to a technical indicator or engineered variable for a given stock at a specific time. Let $\mathscr{F}$ denote the functional space of all possible regression trees, where each $f_k \in \mathscr{F}$ maps an input x_i to a real-valued output. The overall prediction before classification is given by:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathscr{F},$$

where $\hat{y}_i$ is the raw output score for observation i, often referred to as the **margin** or **log-odds**. Each f_k is a regression tree that partitions the feature space and assigns a constant score to each leaf node. The sum of outputs from all K trees yields the model's uncalibrated prediction. For binary classification tasks, this raw output is passed through the logistic sigmoid function:

$$\hat{p}_i = \frac{1}{1 + \exp(-\hat{y}_i)}.$$

This transforms the margin into a probability estimate $\hat{p}_i \in [0, 1]$ that can be interpreted as the model's internal belief about the likelihood that observation i belongs to the positive class. In this application, this is interpreted as the model's estimated probability that a stock's return will exceed the 2% threshold over a 20-day horizon. While this probability does not guarantee an actual return above 2%, it serves as a scoring mechanism for ranking assets by predicted upside potential. To train the model, XGBoost minimizes the following **regularized objective function**:

$$\mathscr{L}(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k),$$

where $l(y_i, \hat{y}_i)$ is the loss function (commonly binary cross-entropy):

$$l(y_i, \hat{y}_i) = -[y_i \log(\hat{p}_i) + (1 - y_i) \log(1 - \hat{p}_i)],$$

and $\Omega(f_k)$ is a regularization term penalizing model complexity:

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} w_j^2.$$

Here, T denotes the number of leaves in the tree f_k, w_j is the score assigned to leaf j while γ and λ are hyperparameters controlling tree sparsity and weight regularization, respectively. At each boosting iteration t, XGBoost adds a new function $f_t(x_i)$ to incrementally improve predictions:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i).$$

During tree construction, XGBoost evaluates potential feature splits by calculating the **information gain**, a measure of how much a split improves the objective function. It does this using both the first-order (gradient) and second-order (Hessian) derivatives of the loss with respect to the predicted margin:

$$\text{Gain} = \frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_R+\lambda} - \frac{(G_L+G_R)^2}{H_L+H_R+\lambda}\right] - \gamma,$$

where G_L and H_L are the sums of gradients and Hessians for the left child node, and G_R and H_R for the right. This formula allows XGBoost to efficiently evaluate split candidates, even in the presence of missing values (GeeksforGeeks 2023). XGBoost is well-suited for financial time series applications due to its:

- **Built-in regularization**, which helps mitigate overfitting in noisy financial data,
- **Soft classification output**, which provides probability-based predictions for ranking assets,
- **Scalability and efficiency**, allowing fast training on large sets of stocks and features, according to Zhang (2022).

In this project, the predicted probabilities from XGBoost are combined with historical return data to compute expected return estimates for use in portfolio optimization.

2.2 Markowitz Mean-Variance Model

The Markowitz mean-variance model, introduced by Harry Markowitz (1952), is a cornerstone of modern portfolio theory. In this model, investment return and risk are quantified by expected return and the variance of returns, respectively. Rational investors aim to either minimize risk for a given expected return or maximize expected return for a given level of risk, balancing these two objectives to maximize expected utility.
The classical formulation of the mean-variance model is a **multi-objective optimization** problem:

$$\begin{aligned} \text{minimize:} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij} \\ \text{maximize:} \quad & \sum_{i=1}^{n} x_i \mu_i \end{aligned}$$

subject to:

$$\begin{aligned} & \sum_{i=1}^{n} x_i = 1, \\ & 0 \le x_i \le 1, \quad \forall i = 1, \dots, n, \end{aligned}$$

where x_i and x_j represent the portfolio weights of assets i and j, σ_{ij} denotes the covariance between asset i and j, and μ_i is the expected return of asset i.

To facilitate practical optimization, the two objectives can be combined into a single **mono-objective formulation** using a risk aversion coefficient $\lambda \in [0, 1]$:

$$\text{minimize:} \quad \lambda \left(\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} \right) - (1 - \lambda) \left(\sum_{i=1}^{n} x_i \mu_i \right)$$

subject to:

$$\sum_{i=1}^{n} x_i = 1,$$
$$0 \leq x_i \leq 1, \quad \forall i = 1, \ldots, n.$$

Here, the risk aversion parameter λ reflects the investor's preference. $\lambda = 0$ represents an extremely risk-seeking investor focused solely on maximizing returns, whereas $\lambda = 1$ corresponds to an extremely risk-averse investor focused solely on minimizing risk. Values of λ between 0 and 1 allow investors to find a balance between return and risk according to their individual preferences.

2.2.1 Black–Litterman Model

The Black-Litterman model, developed by Fischer Black and Robert Litterman at Goldman Sachs, extends the traditional mean–variance framework by incorporating both market equilibrium information and subjective investor views in a consistent Bayesian framework, according to Cabrera (2024). It was originally designed to address several known issues with classical Markowitz optimization, most notably, the sensitivity to input parameters and the tendency to produce extreme, highly concentrated portfolios. The model begins by computing a vector of **implied equilibrium returns** Π based on the assumption that the current market portfolio is mean-variance optimal. These implied returns are inferred through **reverse optimization**, using the formula:

$$\Pi = \lambda \Sigma w_m,$$

where λ is the risk aversion coefficient, Σ is the covariance matrix of asset returns, and w_m is the vector of market capitalization weights. This approach produces a prior estimate of expected returns that is consistent with observed market prices and portfolio structure. Next, the investor's subjective views are introduced. These are encoded in:

- P: a matrix specifying which assets the views relate to and in what combination,
- Q: a vector expressing the expected returns implied by those views,
- Ω: a diagonal covariance matrix representing the uncertainty (or confidence) associated with each view.

The adjusted expected returns are then computed using a Bayesian update:

$$\mathbb{E}(R) = \left[(\tau\Sigma)^{-1} + P^\top \Omega^{-1} P\right]^{-1} \left[(\tau\Sigma)^{-1}\Pi + P^\top \Omega^{-1} Q\right],$$

where τ is a scalar controlling the uncertainty of the prior (Π). The result is a revised return vector $\mathbb{E}(R)$ that blends market consensus with investor conviction. By balancing objective market data with subjective insights, the Black-Litterman model provides more stable and diversified portfolios than classical optimization. In this project, the model is used to integrate machine learning-based predictions into the return vector Q, allowing the portfolio to reflect data-driven views while remaining anchored to market equilibrium.

2.3 Data Description and Preprocessing

The data used in this project can be divided into two main categories: (i) input data for the machine learning model and (ii) input data for the portfolio optimization.

(i) Machine Learning Input Data:
The primary dataset consists of standard OHLC (Open, High, Low, Close) stock price data obtained from StockAnalysis (StockAnalysis 2025). To enhance predictive performance, various lagged variables and technical indicators such as EMA, RSI, and Bollinger Bands (stockindicators 2025) were engineered from the OHLC data. These features serve as input variables for the XGBoost classification model, which aims to predict the probability of achieving a stock return greater than 2% over a 20-day horizon. Feature selection was performed using SHAP (SHapley Additive exPlanations) analysis (GeeksforGeeks 2025b), which evaluates the contribution of each feature to the model's predictions. For each stock, the top-ranked features based on SHAP values were selected as the final set of input variables. To visualize the feature engineering process, plots showing the stock's close price along with selected technical indicators were produced, providing insights into the relationship between technical signals and price movements.

(ii) Portfolio Optimization Input Data:
Following the machine learning stage, a separate CSV file was constructed containing summarized model outputs for each stock. This table includes key information such as the predicted probability of achieving a >2% return, the investor verdict (BUY/HOLD/SELL), and evaluation metrics (see Table 5). A filtering function was applied to select stocks with both the highest model prediction quality and the highest predicted probabilities. Thus, the data feeding into the portfolio optimization is directly derived from the machine learning results, ensuring a dynamic and data-driven investment selection process. Probabilistic predictions from the XGBoost model are not only used for classification, but also transformed into expected returns. The predicted probabilities are then passed into the optimization models to guide stock selection and allocation. This ensures a fully integrated data-driven investment framework, where machine learning outputs directly inform financial decision-making.

2.4 Machine Learning for Stock Predictions

The machine learning stage aims to predict the probability that a stock will achieve a return greater than 2% over a 20-day horizon. An XGBoost classification model is employed for this purpose, using features derived from technical indicators and lagged price movements. A range of technical indicators such as MACD, RSI and Bollinger Bands (stockindicators 2025) are computed from OHLC data (StockAnalysis 2025), together with volatility and trend-based features. Figure 1 illustrates an example of a stock's close price alongside selected technical indicators.

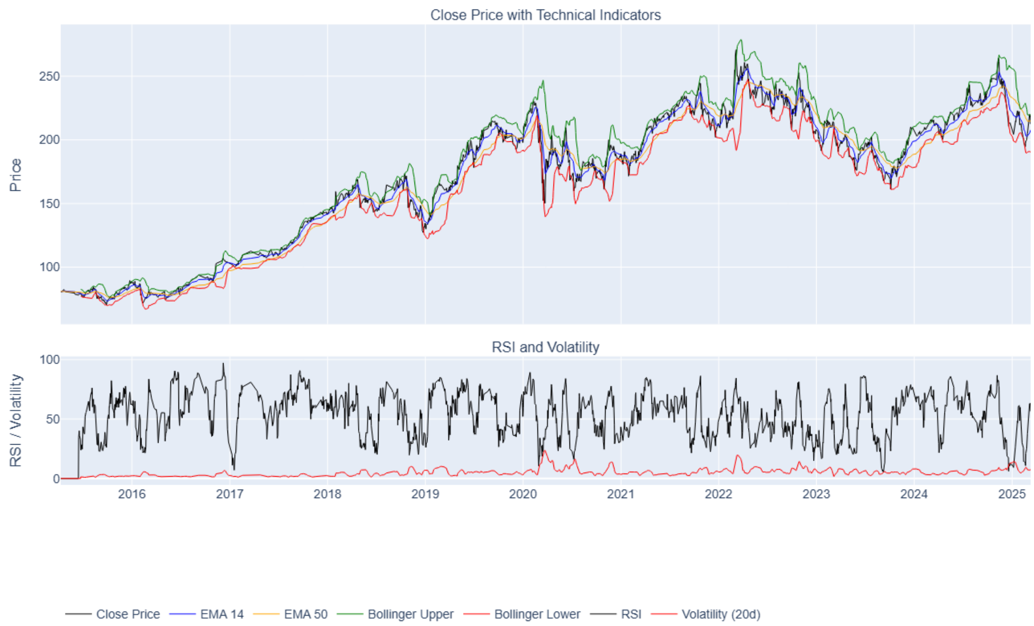

Figure 1: Close Price with Technical Indicators for LHX

The dataset is split chronologically into training (70%), validation (15%), and test (15%) sets to preserve the time structure of financial data. All features are scaled to the range $[-1, 1]$ using MinMaxScaler. To address class imbalance, different resampling methods are evaluated, including SMOTE (GeeksforGeeks 2024) and random undersampling, selecting the approach that produces the best validation F1 score. Feature selection is conducted using SHAP analysis, identifying the most important variables driving the model's predictions. The top-ranked features are retained for model training to enhance both performance and interpretability.

Hyperparameter tuning is performed using Optuna, as described by Loomis (2020), with time-series cross-validation, optimizing for the ROC AUC score. In addition, an exponential decay weighting scheme is applied during training to prioritize more recent data points. The predictive performance of the model is evaluated based on the ROC AUC and Brier Score, as explained by Gongting (2019), the PR AUC introduced by deepchecks (2025) and Log Loss described by KoshurAI (2024). Furthermore, the predicted probability distribution and the ROC curve is visualized to evaluate the calibration and discriminative capacity of the model. Finally, to benchmark performance, the XGBoost model is compared against two baseline models: a **Dummy Classifier**, explained by Baladram (2024) and a **Logistic Regression Model** (GeeksforGeeks 2025a). The Dummy Classifier serves as a naive benchmark by making predictions based on simple rules such as always choosing the most frequent class or making random guesses. Logistic Regression is a standard linear classification model that estimates the probability of class membership using a weighted combination of input features passed through a sigmoid function. These baselines help assess whether the XGBoost model provides meaningful improvements over simpler or non-informative approaches. The resulting probability outputs from each model form the basis for subsequent portfolio optimization.

2.5 Portfolio Optimization

The optimization was conducted on a filtered subset of stocks that were selected based on machine learning predictions. Specifically, stocks were required to have a predicted probability of exceeding a 2% return over a 20−day horizon above 30%, and a ROC AUC score greater than 0.55. This ensured that only the most promising assets were included in the portfolio construction phase, based on predictive returns and model performance. Expected returns (μ_i) for each asset were computed by blending two information sources: the predicted probability from the XGBoost classifier and the average historical return over a 20-day window. This hybrid approach ensures that the forecasted returns incorporate both forward-looking model insights and past performance. The formula used is:

$$\mu_i = (w \cdot p_i \cdot 0.02 + (1 - w) \cdot r_i)^{\frac{252}{20}} - 1$$

Here, w is a weighting parameter (set to 0.6) controlling the influence of the machine learning model, p_i is the predicted probability that the stock will exceed a 2% return over the 20-day horizon, and r_i is the historical mean return over the same period. The term $p_i \cdot 0.02$ represents the expected value of a 2% return conditioned on the model's probability estimate. The expression is then annualized by raising the blended return to the power of 252/20, corresponding to the number of trading periods in a year. This transformation ensures consistency with the annualized covariance matrix used in portfolio optimization.

Monte Carlo simulations were employed to explore a wide range of potential portfolio configurations under different asset combinations and weightings. In each simulation run, a random subset of k assets was selected from the full pool of eligible stocks, where $k \in \{6, 7, 8, 9, 10\}$. For the selected k assets, a set of random portfolio weights was generated from a uniform distribution and then normalized to ensure that the weights summed to 1, representing a fully invested, long-only portfolio. This process was repeated thousands of times to generate a diverse set of feasible portfolios across different sizes and compositions. Each simulated portfolio was evaluated based on its expected return, volatility, and Sharpe ratio, forming the basis for identifying optimal portfolios. In the Markowitz approach, portfolios with the highest Sharpe ratios, minimum volatility and maximum expected return were retained as optimal for each cardinality level. For the Black-Litterman model, market-implied returns (Π) were first derived from market capitalization. Machine learning predictions were incorporated into the Black-Litterman framework by treating them as subjective investor views. Specifically, the predicted expected returns from the XGBoost model were used to construct the Q vector of views, while the P matrix was set to the identity matrix to reflect views on individual assets. To account for uncertainty in the model predictions, the diagonal covariance matrix Ω, which expresses confidence in each view, was scaled according to the model's log loss, assigning lower confidence to less reliable predictions. The scalar τ was used to control the influence of the market equilibrium prior relative to the subjective views, following the standard shrinkage formulation. The resulting posterior expected returns μ_{BL}, a blend of market equilibrium and model-informed views, were then used as inputs to the same Monte Carlo simulation and optimization routine as in the Markowitz approach. This process yielded a collection of optimized portfolios under different assumptions and modeling philosophies.

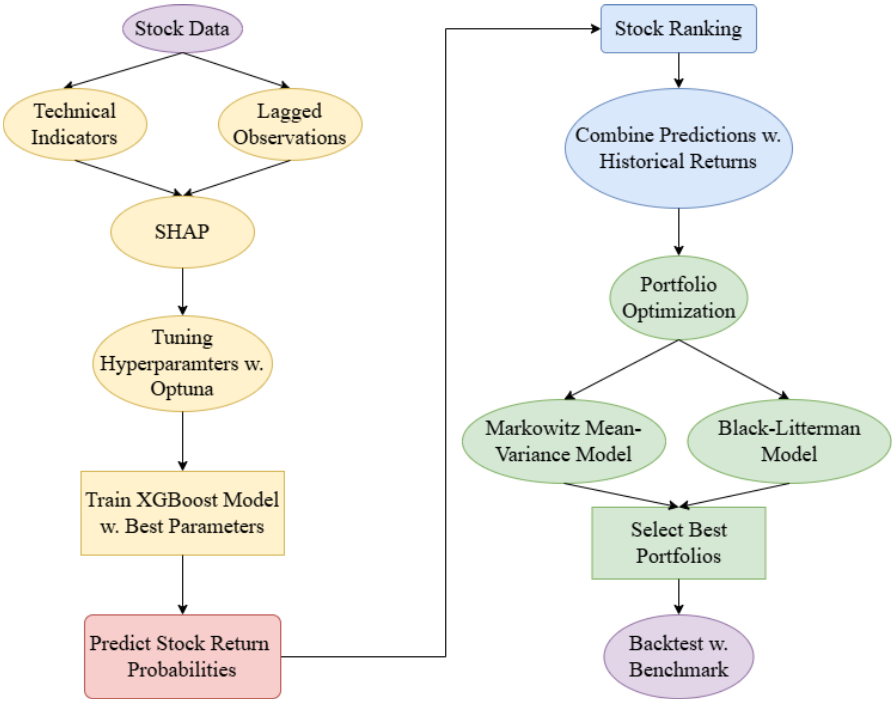

Figure 2: Flowchart for Project Structure

3 Analysis of Results

This section presents and analyzes the results of the machine learning model and portfolio optimization strategies developed in this thesis. It begins by evaluating classification performance through ROC curves and model comparison tables across different stocks, highlighting strengths and limitations of the predictive models. Next, the distribution of predicted probabilities is examined to assess model confidence. The second part focuses on portfolio construction and evaluation, showcasing the composition and performance of optimal portfolios selected from the efficient frontier. To conclude, out-of-sample backtests are conducted to assess real-world robustness, comparing all optimized portfolios against the SPY benchmark using key performance metrics.

3.1 Evaluation of Machine Learning Model

3.1.1 Feature Interpretability via SHAP

While SHAP was used in the feature selection process, it is also valuable for interpreting the model's behavior after training. SHAP provides both local (per prediction) and global (overall) feature importance scores by attributing each prediction to individual input variables in a fair and consistent manner. This enables a deeper understanding of which features drive the model's decisions and whether these align with known financial principles. For example, if the model relies heavily on technical indicators like EMA, MACD, or Bollinger Bands, tools commonly used by traders, it suggests that the model is capturing relevant market patterns. Conversely, if the model places high importance on noisy or arbitrary features, it may indicate overfitting or lack of robustness. Thus, SHAP analysis helps evaluate both the interpretability and credibility of the model in a financial context. Figure 3 shows the 15 most influential features identified by SHAP across all training data. As expected, EMA-based signals, MACD and Bollinger Bands appear as dominant drivers of predictions.

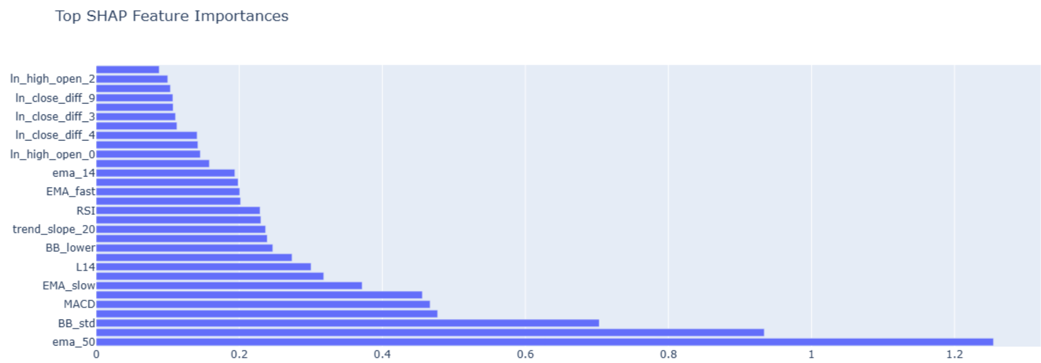

Figure 3: SHAP analysis for LHX

3.1.2 Performance and Limitations

The performance of the XGBoost model varies considerably across individual stocks. This is expected, as stocks differ in their underlying dynamics, volatility structures, and sensitivity to market regimes or external shocks. For instance, stocks in stable sectors may exhibit more predictable trends, while others, especially those with irregular, flat, or highly volatile return patterns, pose greater challenges for models based solely on historical price-based features. Since the model does not incorporate macroeconomic, fundamental, or news-based data, it is more sensitive to the inherent noise or irregularity in a stock's price history. These differences contribute to the observed variation in prediction accuracy across assets. To illustrate these performance differences, two representative stocks are examined: one where the model demonstrates strong predictive power, and another where the model has somewhat weaker performance. These cases reflect typical strengths and weaknesses of using machine learning for stock prediction. The stocks in question are: **LHX (Aerospace & Defense):** A stock with sharp trends and frequent technical signals, where the model achieves consistent and confident probability estimates. **NVDA (Technology):** A stock with a continuous and lower price for many years- until a drastic jump in price that reflects the company‘s success in recent years, where predictions appear less calibrated and performance is very poor.

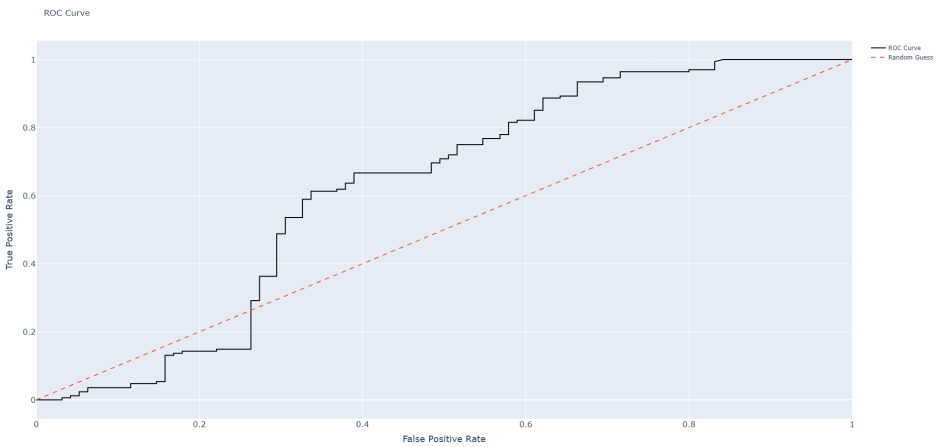

Figure 4: ROC curve for LHX

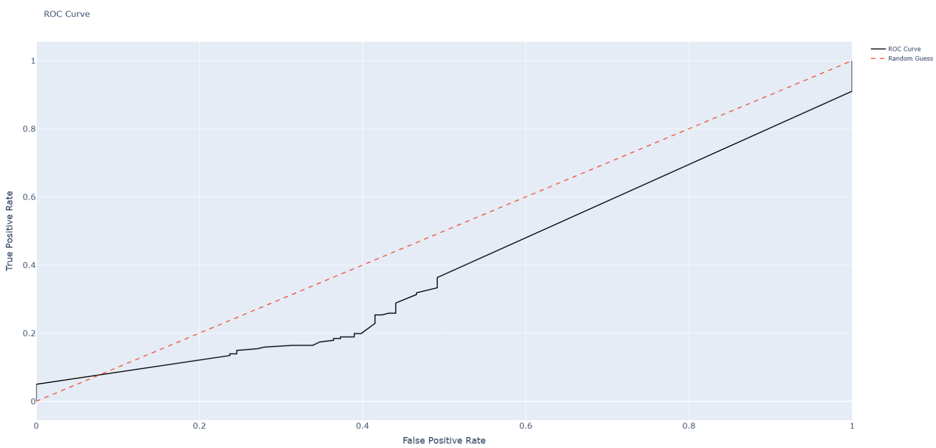

Figure 5: ROC curve for NVDA

Figures 4 and 5 illustrate the ROC curves for two different stocks, highlighting variation in model performance. In these plots, a curve closer to the top-left corner indicates better discrimination between the classes, and a larger area under the curve (AUC) reflects stronger predictive power. The contrast between the two figures shows how classification quality can vary across assets, even when using the same model architecture and features.

3.1.3 Comparative Results Across Models

Table 1: Comparison of model performance for LHX and NVDA

Model	ROC AUC	PR AUC	Log Loss	Brier Score
LHX				
XGBoost	0.64	0.68	0.62	0.21
Logistic Regression	0.58	0.66	0.64	0.22
Dummy Classifier	0.50	0.82	13.01	0.36
NVDA				
XGBoost	0.37	0.56	1.11	0.42
Logistic Regression	0.69	0.79	3.05	0.52
Dummy Classifier	0.50	0.82	13.33	0.37

Table 1 compares the predictive performance of the XGBoost classifier against two baseline models: a Logistic Regression and a Dummy Classifier, for the two stocks LHX and NVDA. For LHX, the XGBoost model achieves the best ROC AUC and Log Loss values, indicating a reasonably well-calibrated and discriminative model. In contrast, the Dummy Classifier fails to provide meaningful predictions. Logistic Regression performs moderately, but is consistently outperformed by XGBoost on this stock.

Interestingly, for NVDA, the performance ranking reverses. Here, Logistic Regression achieves a significantly higher ROC AUC than XGBoost, which struggles to discriminate between classes. This likely reflects the model's difficulty in capturing NVDA's recent structural trend shift using only historical price-based features. The poor calibration and weak separation of predicted probabilities are further indicated by the relatively high Log Loss and Brier Score for XGBoost. These results illustrate that although the XGBoost model is more powerful in many cases, it does not universally outperform simpler models like Logistic Regression. Model performance is thus context-dependent and highlights the importance of robustness checks before portfolio application. Although one can use the predictions for the better-performing model for different stocks, this project has focused on one machine learning model only, the XGBoost. The full analysis of all stocks used in this project can be viewed in Table 5 in Section 7.

3.1.4 Probability Distributions and Confidence

Beyond traditional performance metrics, it is also important to examine how predicted probabilities are distributed, i.e. how confident the model is in classifying each instance. In the context of binary stock return prediction, probabilities close to 0 or 1 indicate stronger classification confidence, whereas values near 0.5 reflect model indecision or weak signal strength. Figure 6 and Figure 7 show the distribution of predicted probabilities for the stocks LHX and NVDA, respectively. For LHX, the model displays a skewed distribution leaning towards higher probability values (above 0.6), suggesting stronger separation between classes and more decisive predictions. This pattern typically reflects higher model confidence and is often associated with better class discrimination when supported by complementary metrics like ROC AUC and log loss.

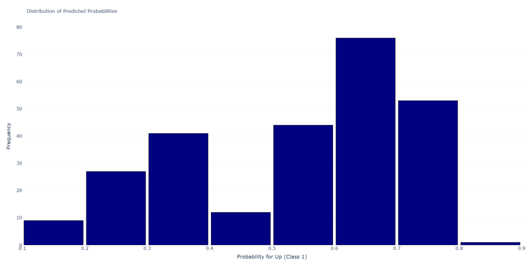

Figure 6: Distribution of predicted probabilities for LHX

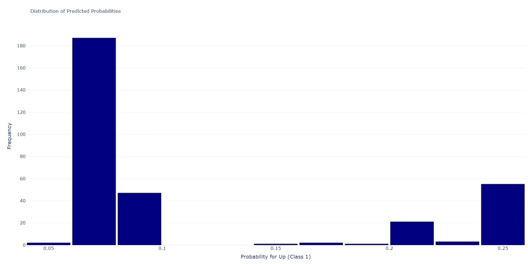

Figure 7: Distribution of predicted probabilities for NVDA

In contrast, the NVDA distribution is narrowly concentrated between 0.05 and 0.1, indicating that the model consistently assigns low probabilities to the positive class. This pattern suggests low predictive confidence and signal quality for this stock, and is consistent with its weaker ROC AUC performance. These examples illustrate how probability distributions can provide valuable diagnostic insight beyond accuracy metrics. Stocks where the model outputs are uncertain, i.e. probability distributions are flat or clustered around low values, may lead to suboptimal decisions if treated equally in the portfolio construction step. Thus, such distributions help assess when and where the model's predictions are most actionable. These examples show that probability distributions can offer additional diagnostic insight alongside traditional metrics. For NVDA, the narrow concentration of low predicted probabilities aligns with its weaker ROC AUC and higher log loss, indicating poor class separation. In contrast, LHX shows a broader distribution skewed toward the positive class, which is consistent with stronger model performance. While these observations are based on a limited set of stocks, they suggest that examining distribution shape can help assess predictive confidence and uncertainty-important considerations in portfolio construction.

3.2 Portfolio Optimization Results

3.2.1 Overview of Selected Assets

Based on the filtering criteria, a total of $\mathbf{n = 41}$ **stocks** were selected for portfolio optimization. These assets were chosen based on a combination of machine learning-based probability predictions, risk metrics, and investor verdicts. The filtering process ensured a preliminary level of diversification, with the selected stocks spanning multiple sectors and some geographic regions. To illustrate the distribution of the selected assets, Figure 8 presents a sector-wise breakdown of the final stock pool. Each bar represents the number of stocks within a given sector.

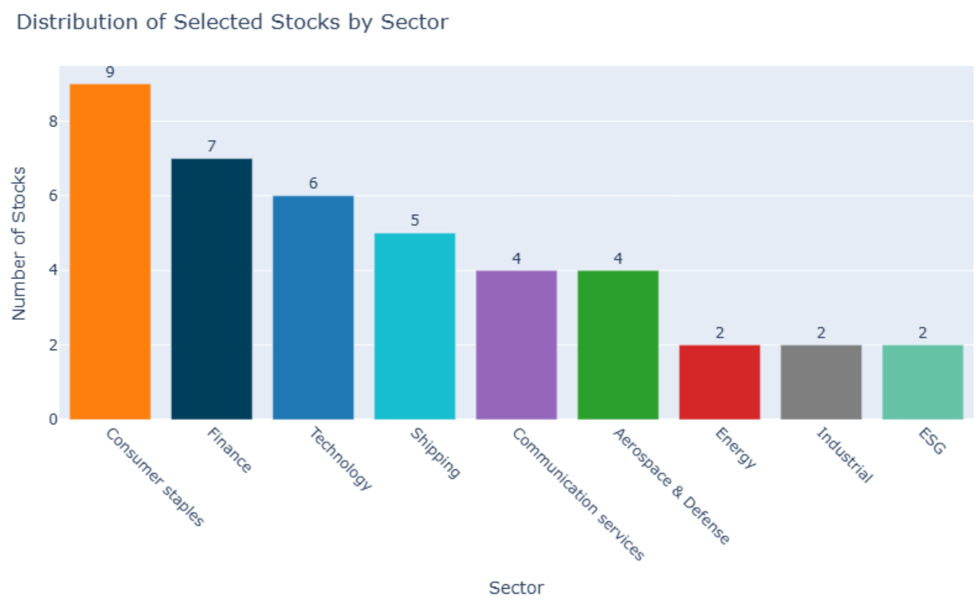

Figure 8: Sector Distribution

3.2.2 Efficient Frontier Results

The **efficient frontier**, described by Adam (2023) is a fundamental concept in portfolio theory that represents the set of optimal portfolios offering the highest expected return for a given level of risk, or equivalently, the lowest risk for a given level of expected return. Portfolios lying on the efficient frontier are considered *Pareto optimal*, meaning no other portfolio exists that improves one objective (return or risk) without worsening the other. In this project, the efficient frontier was constructed using Monte Carlo simulation, where $100,000$ portfolios were randomly generated for each model (MVO and BL). Each simulated portfolio is defined by a unique combination of asset weights across the filtered stock universe. For each portfolio, the expected return μ_i is computed as the weighted average of the individual asset expected returns. These asset-level expectations differ depending on the model used:

- For the Markowitz model, expected returns μ_i are estimated as a weighted combination of historical average returns and predicted probability-adjusted returns from the XGBoost classifier. Specifically, predicted probabilities of exceeding a 2% return are scaled and annualized, then blended with historical means to form forward-looking estimates.

- For the Black-Litterman model, expected returns are derived using a Bayesian adjustment of the market-implied equilibrium returns Π, where investor views are informed by the same machine learning predictions. The degree of influence from machine learning forecasts is controlled through the parameters τ and Ω.

Volatility for each portfolio is computed using the covariance matrix of historical returns, assuming constant correlations across assets. Together, each point on the frontier represents a feasible portfolio, with its own risk-return characteristics shaped by the interplay between model-based predictions, diversification, and asset covariance. This simulation-based approach allows visualizing how machine learning-enhanced expectations shift the efficient frontier relative to traditional approaches. Figure 9 and Figure 10 below illustrate the respective frontiers.

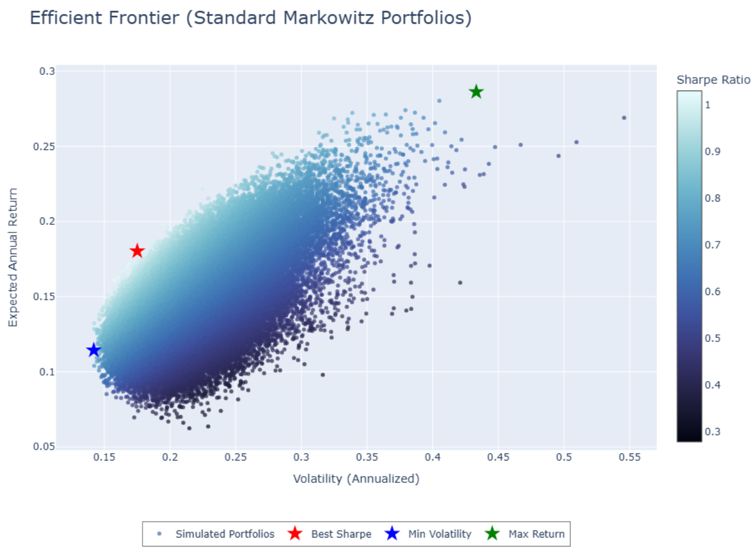

Figure 9: Efficient frontier for Markowitz Model

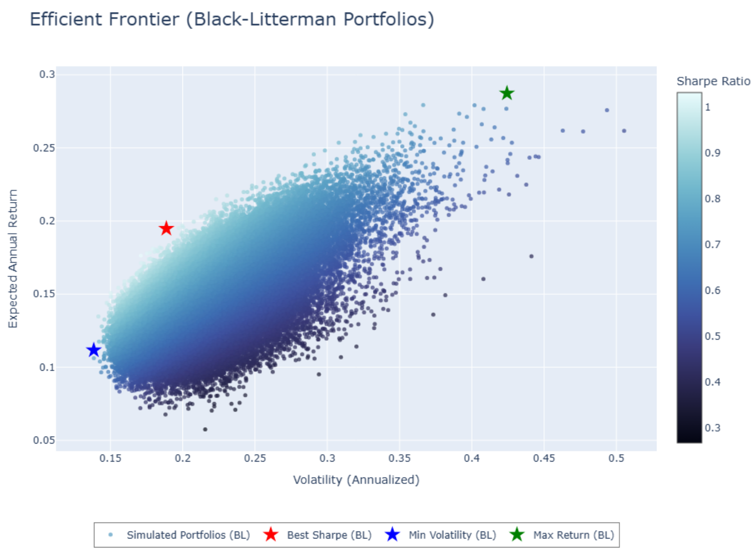

Figure 10: Efficient frontier for Black-Litterman Model

From the cloud of optimal portfolios, three **highlighted portfolios** are extracted and marked with distinct stars in each of the plots above. The *red star* represents the portfolio with the highest **Sharpe Ratio**, offering the best trade-off between risk and return. The *blue star* marks the **minimum volatility portfolio**, ideal for risk-averse investors. Lastly, the *green star* shows the **maximum return portfolio**, which typically carries higher risk but may appeal to return-maximizing strategies. The three highlighted portfolios (red, blue, and green stars) correspond to the highest Sharpe ratio, minimum volatility, and maximum return portfolios, respectively. Importantly, these were not computed separately, but were extracted directly from the 100,000 simulated portfolios, ensuring they lie within the actual efficient frontier generated by the Monte Carlo simulation. Their positioning reflects naturally occurring optima within the simulated portfolio universe. In Figure 9, the efficient frontier is constructed using historical returns adjusted with machine learning forecasts. The frontier has a relatively wide dispersion, with the best Sharpe portfolio located at a moderate risk-return balance, and the maximum return portfolio exhibiting both high expected return and high volatility.

In contrast, Figure 10 demonstrates a slightly more compact and upward-shifted frontier, indicating improved risk-adjusted returns. The best Sharpe portfolio under the Black-Litterman approach appears to achieve a higher expected return for a similar level of risk compared to the Markowitz equivalent. Moreover, the min-volatility and max-return portfolios are positioned differently, reflecting the influence of the investor‘s views and market equilibrium inputs in the Black-Litterman framework. Overall, the Black-Litterman model shows a slightly more favorable efficient frontier curvature, suggesting better portfolio diversification and alignment with both data-driven and market-implied expectations.

To provide further insight into the portfolios highlighted in the efficient frontier plots, the following tables detail the asset weights, sectors, and regions of two key portfolios from the simulation: the **Markowitz Maximum Return** Table 2 and the **Black-Litterman Best Sharpe Ratio** Table 3 portfolios. These two portfolios were selected to illustrate contrasting investment strategies. The Markowitz Maximum Return portfolio represents an aggressive, high-risk approach focused on maximizing expected return, while the Black-Litterman Best Sharpe Ratio portfolio demonstrates a more balanced and diversified allocation with superior risk-adjusted performance. Presenting both offers a clear view of the trade-off between return potential and risk control across the two modeling frameworks. The remaining portfolios are listed in the appendix sec 7.

Portfolio Statistics – Markowitz Maximum Return

Expected Return:	28.62%	Sharpe Ratio:	0.6607
Volatility:	43.32%	VaR (95%):	42.64%

Stock	**Weight**	**Sector**	**Region**
NEE	0.73%	ESG	US
MATX	19.01%	Shipping	US
MMM	0.55%	Industrial	US
AMD	28.07%	Technology	US
ENPH	41.84%	ESG	US
EOG	9.81%	Energy	US

Table 2: Asset allocation for the Markowitz Maximum Return (green star in Figure 9).

Portfolio Statistics – Black-Litterman Best Sharpe Ratio

Expected Return:	19.88%	Sharpe Ratio:	1.0536
Volatility:	18.87%	VaR (95%):	11.16%

Stock	Weight	Sector	Region
NOC	32.38%	Aerospace & Defense	US
NEE	14.75%	ESG	US
HSBC	0.48%	Finance	EU
AMD	12.75%	Technology	US
DSDVF	15.16%	Shipping	EU
AAPL	24.46%	Technology	US

Table 3: Asset allocation for the Black-Litterman Best Sharpe Portfolio (red star in Figure 10).

3.2.3 Backtesting Optimal Portfolios

To evaluate the real-world performance of the optimized portfolios, backtesting procedure was conducted over the out-of-sample period from **Januray 2021 to March 2025**. Each of the six optimal portfolios, three from the Markowitz model and three from the Black-Litterman model, was tested using fixed weigths derived from the efficient frontier simulation. The goal of this backtest is to assess how well the portfolios would have performed historically in terms of cumulative growth, risk, and return characteristics. The analysis includes the portfolio with the highest Sharpe ratio, the minimum volatility portfolio, and the maximum return portfolio from each model. Given the diversified nature of the constructed portfolios across multiple sectors and regions, the benchmark used for comparison is the *SPDR S&P 500 ETF Trust* ***(SPY)***. SPY represents a broad-market index and provides a more appropriate reference point than tech-heavy benchmarks such as **QQQ**, which is excluded from the plots to maintain clarity and sectoral relevance.

Figure 11: Minimum volatility-portfolios vs. SPY

Figure 12: Best Sharpe ratio-portfolios vs. SPY

Figure 11, Figure 12, and Figure 13 show the out-of-sample value development of the minimum volatility, best Sharpe ratio- and maximum return portfolios respectively. Each portfolio is compared against the SPY benchmark to assess performance stability and trend behavior over time. To systematically evaluate performance, four core metrics are used: **CAGR**, explained by Fernando (2024a) reflects the annualized return assuming continuous compounding over the backtest period. It captures long-term growth and is central to assessing investment attractiveness. **Annualized Volatility** measures the standard deviation of portfolio returns scaled to a yearly basis. It serves as a proxy for risk, with higher values indicating greater fluctuations in value. **Sharpe Ratio**, according to Fernando (2024b) quantifies risk-adjusted return by dividing excess return by volatility. A higher Sharpe ratio implies more return per unit of risk, making it especially relevant when comparing portfolios of varying risk levels. **Maximum Drawdown** captures the largest observed peak-to-trough decline in portfolio value during the test period. It reflects downside risk and potential for capital loss, offering insight into worst-case performance under stress. These metrics, summarized in Table 4, enable a comprehensive assessment of return potential, stability, and risk for each portfolio relative to the SPY benchmark.

As seen in Figure 11, the minimum volatility portfolios exhibit relatively smooth value trajectories with fewer sharp drawdowns compared to the SPY benchmark. This aligns with their optimization objective and highlights their suitability for risk-averse investors. Similarly, the best Sharpe ratio portfolios in Figure 12 demonstrate consistent upward trends while avoiding extreme volatility, indicating strong risk-adjusted performance. The Black-Litterman variant appears to track a steadier growth path with less deviation than its Markowitz counterpart, suggesting improved stability through integration of market equilibrium views. These visual patterns will be further quantified in the next section. In contrast, the maximum return portfolios (see Figure 13) exhibit much higher volatility and significant fluctuations throughout the test period. While these portfolios achieve the highest expected return on paper, their unstable value paths and deep drawdowns underscore the trade-off between aggressive return-seeking and real-world performance reliability. This highlights the importance of including risk-aware metrics in performance evaluation, not just return.

Figure 13: Maximum return-portfolios vs. SPY

Portfolio	CAGR	Volatility	Sharpe Ratio	Max Drawdown
Markowitz (Best Sharpe)	12.38%	17.83%	0.74	-24.10%
Black-Litterman (Best Sharpe)	12.39%	18.59%	0.72	-20.43%
Markowitz (Min Volatility)	10.35%	13.56%	0.79	-21.54%
Black-Litterman (Min Volatility)	10.51%	13.59%	0.80	-21.16%
Markowitz (Max Return)	2.89%	38.81%	0.27	-38.24%
Black-Litterman (Max Return)	3.31%	38.75%	0.28	-40.09%
SPY Benchmark	10.48%	16.49%	0.69	-25.36%

Table 4: Summary of backtest performance metrics (2021–2025) for the six optimized portfolios compared to the SPY benchmark.

4 Discussion

The results of the portfolio optimization and subsequent backtesting offer several key insights into the relative performance and robustness of the Markowitz and Black-Litterman frameworks. While both models were able to produce portfolios with favorable performance compared to the SPY benchmark, some notable differences emerged with respect to risk-adjusted returns, volatility control, and practical stability. First, the Black-Litterman portfolios outperformed their Markowitz counterparts on risk-adjusted metrics, particularly in the case of the minimum volatility and best Sharpe ratio portfolios. The results are quite even in terms of Sharpe ratio, but the Black-Litterman approach seems to outperform the Markowitz model for CAGR and max drawdown in most cases. This suggests that integrating market equilibrium views with investor-specific opinions, in this case derived from machine learning-based predictions, can lead to more stable and efficient allocations. The ability of Black-Litterman to smooth out extreme input sensitivities, a well-known limitation of classical mean-variance optimization, appears to have had a tangible effect on real-world backtest performance.

Second, the maximum return portfolios in both models exhibited high volatility and underperformed significantly in terms of cumulative return, as seen in Table 4. This reinforces the classic trade-off between risk and return: portfolios designed purely to maximize expected return tend to suffer from poor downside control and large drawdowns. These results highlight the importance of considering volatility and risk preferences explicitly during the optimization phase, rather than pursuing return in isolation.

Third, the use of fixed portfolio weights throughout the backtest implies that the portfolios were not rebalanced during the out-of-sample period. While this approach allows for a clean comparison of model performance, it does not account for changing market conditions or predictive signals over time. A rolling or dynamically updated strategy might yield improved results, albeit at the cost of increased complexity and transaction considerations. Finally, while the SPY ETF served as a reasonable benchmark due to its diversified sector exposure, it is worth noting that the optimized portfolios are not intended to track any specific index. Rather, they are constructed to maximize return or Sharpe ratio given a filtered universe of stocks. Therefore, outperformance or underperformance relative to SPY should be interpreted in light of these design objectives. Overall, the analysis supports the conclusion that combining data-driven predictive signals with structured portfolio models such as Black-Litterman can yield meaningful improvements in out-of-sample performance. At the same time, the study underscores the importance of balancing return objectives with risk control, and of carefully selecting evaluation metrics to reflect the portfolio's intended purpose.

The results of the portfolio optimization and backtesting provide meaningful insights into the interaction between machine learning-based signals and classical financial optimization models. While the quantitative metrics such as Sharpe ratio and volatility provide a strong baseline for model comparison, it is equally important to examine the qualitative characteristics of the resulting portfolios, particularly asset allocation, sector exposure, and geographic spread. Looking at the *Black-Litterman Best Sharpe Ratio* portfolio (see Table 3), it is noteworthy that the allocation is fairly balanced across sectors such as Aerospace & Defense (NOC), ESG (NEE), Technology (AAPL, AMD), Finance (HSBC), and Shipping (DSDVF). This blend reflects a structurally diversified portfolio, one that aligns with risk-adjusted return maximization under market equilibrium views and predictive adjustments. The presence of both defensive sectors (Aerospace, Finance) and growth-oriented assets (Technology, ESG) suggests robustness to varying macroeconomic conditions. In contrast, the *Markowitz Maximum Return* portfolio (see Table 2) is heavily concentrated in a few high-beta assets, such as ENPH and AMD, which explains its high expected return but also extreme volatility (over 40%) from Table 4.

The integration of XGBoost-based probability estimates into expected return calculations proved valuable. Stocks like AMD and NEE consistently received high weights across several portfolios, driven by strong signals from the machine learning model. This demonstrates that well-tuned classification models can provide directional insights that influence asset selection beyond pure historical averages. However, it's also clear that not all machine learning-favored stocks are included in the best-performing portfolios, reflecting the trade-off between predicted return and portfolio-level risk (see Section 7 for analysis of all stocks). At the same time, it is worth noting that machine learning is not a concrete strategy for predicting stock returns. However, machine learning might learn consecutive patterns for individual stocks that fluctuate in a repetitive manner.

From a macroeconomic-and-geopolitical standpoint, several of the top-weighted stocks across portfolios align with global megatrends. For example: **Aerospace & Defense** (e.g., NOC, RTX) has likely benefited from rising geopolitical tensions and increased defense spending across the US and Europe. **Technology** firms (AMD, MSFT, AAPL) continue to be favored due to strong innovation pipelines, AI demand, and global digitalization. **ESG-focused** companies (NEE, ENPH) have high exposure to clean energy, aligning with long-term policy and investment flows. Portfolios with a broader global and sectoral exposure, such as the Black-Litterman Best Sharpe (see Table 3) or the Black-Litterman Minimum Volatility (see Section 7), may be better suited to structural shifts, compared to those that are concentrated in volatile or cyclical sectors like pure energy or shipping. Finally, the contrast between the maximum return and best Sharpe portfolios underscores the critical difference between chasing raw returns and achieving balanced performance. While the maximum return portfolios recorded the highest expected returns, their excessive volatility and poor backtest performance (Sharpe $\approx$ 0.27–0.28) demonstrate the pitfalls of unbalanced exposure. Conversely, minimum volatility portfolios provided steadier performance with more attractive drawdown profiles, albeit at lower returns. For investors or analysts, these findings suggest that incorporating forward-looking views, via machine learning and combining them with equilibrium-based frameworks like Black-Litterman may yield portfolios that are not only quantitatively optimal but also strategically aligned with real-world developments. Furthermore, sector- and region-aware allocations provide an essential buffer against sudden market shifts or concentration risk.

5 Conclusion

This thesis explored the integration of machine learning and portfolio optimization to improve equity investment strategies. A probabilistic XGBoost classifier was used to forecast the likelihood of short-term positive stock returns, and these predictions were incorporated into both the Markowitz and Black-Litterman frameworks. The results show that portfolios optimized under the Black-Litterman model generally achieved high Sharpe ratios and more stable performance compared to the Markowitz equivalents. Minimum volatility and Sharpe-optimal portfolios consistently outperformed their maximum-return counterparts in terms of drawdown and overall robustness. These findings support the idea that combining forward-looking machine learning signals with structured optimization frameworks can lead to more balanced and resilient investment portfolios. This hybrid approach represents a promising direction for data-driven asset management.

6 References

Adam (2023). *Solving for the Efficient Frontier in Stock Portfolios*. Medium. URL: https://medium.com/call-for-atlas/solving-for-the-efficient-frontier-in-stock-portfolios-6d8d4e7a6f89 (visited on 04/04/2025).

Baladram, Samy (2024). *Dummy Classifier Explained: A Visual Guide with Code Examples for Beginners*. towardsdatasciece. URL: https://towardsdatascience.com/dummy-classifier-explained-a-visual-guide-with-code-examples-for-beginners-009ff95fc86e/ (visited on 22/03/2025).

Berk, Jonathan and DeMarzo, Peter (2020). *Corporate Finance*. 5th, Global Edition. Pearson. ISBN: 9781292304151.

Cabrera, Juan (2024). *Black-Litterman Model*. FINANCIALEDGE. URL: https://www.fe.training/free-resources/portfolio-management/black-litterman-model/ (visited on 02/04/2025).

deepchecks (2025). *What is Precision-Recall AUC?* deepchecks. URL: https://www.deepchecks.com/glossary/pr-auc/ (visited on 20/03/2025).

Fernando, Jason (2024a). *Compound Annual Growth Rate (CAGR) Formula and Calculation*. Investopedia. URL: https://www.investopedia.com/terms/c/cagr.asp (visited on 07/04/2025).

— (2024b). *Sharpe Ratio: Definition, Formula, and Examples*. Investopedia. URL: https://www.investopedia.com/terms/s/sharperatio.asp (visited on 25/03/2025).

GeeksforGeeks (2023). *XGBoost*. GeeksForGeeks. URL: https://www.geeksforgeeks.org/xgboost/ (visited on 15/03/2025).

— (2024). *SMOTE for Imbalanced Classification with Python*. GeeksForGeeks. URL: https://www.geeksforgeeks.org/smote-for-imbalanced-classification-with-python/ (visited on 15/03/2025).

— (2025a). *Logistic Regression in Machine Learning*. GeeksForGeeks. URL: https://www.geeksforgeeks.org/understanding-logistic-regression/ (visited on 22/03/2025).

— (2025b). *SHAP: A Comprehensive Guide to SHapley Additive exPlanations*. GeeksForGeeks. URL: https://www.geeksforgeeks.org/shap-a-comprehensive-guide-to-shapley-additive-explanations/ (visited on 15/03/2025).

Gongting, Tina (2019). *Understanding ROC AUC: Pros and Cons. Why is Bier Score a Great Supplement?* Medium. URL: https://medium.com/@penggongting/understanding-roc-auc-pros-and-cons-why-is-bier-score-a-great-supplement-c7a0c976b679 (visited on 20/03/2025).

Hastie, Trevor, Tibshirani, Robert and Friedman, Jerome (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer Series in Statistics. Springer New York.

Kazeem, Oladimeji (Sept. 2023). 'PORTOLIO OPTIMIZATION USING MACHINE LEARNING'. PhD thesis. URL: https://www.researchgate.net/publication/374084001_PORTOLIO_OPTIMIZATION_USING_MACHINE_LEARNING (visited on 25/03/2025).

KoshurAI (2024). *Understanding Log Loss: A Comprehensive Guide with Code Examples*. Medium. URL: https://koshurai.medium.com/understanding-log-loss-a-comprehensive-guide-with-code-examples-c79cf5411426 (visited on 20/03/2025).

Loomis, Crissman (2020). *Using Optuna to Optimize XGBoost Hyperparameters*. Medium. URL: https://medium.com/optuna/using-optuna-to-optimize-xgboost-hyperparameters-63bfcdfd3407 (visited on 15/03/2025).

Markowitz, Harry (1952). 'Portfolio Selection'. In: *The Journal of Finance* 7.1, pp. 77–91. ISSN: 00221082, 15406261. URL: http://www.jstor.org/stable/2975974 (visited on 09/05/2025).
OpenAI (2021). *UiO GPT-3.5: Large Language Model*. Open AI and University of Oslo. URL: https://gpt.uio.no.
Pliska, Stanley R. (1997). *Introduction to Mathematical Finance*. Blackwell Publishing.
StockAnalysis (2025). StockAnalysis. URL: https://stockanalysis.com/ (visited on 05/03/2025).
stockindicators (2025). *Indicators and Overlays*. stock indicators. URL: https://python.stockindicators.dev/indicators/ (visited on 10/03/2025).
Zhang, Yifan (2022). 'Stock Price Prediction Method Based on XGboost Algorithm'. In: *Proceedings of the 2022 International Conference on Bigdata Blockchain and Economy Management (ICBBEM 2022)*. Atlantis Press, pp. 595–603. ISBN: 978-94-6463-030-5. DOI: 10.2991/978-94-6463-030-5_60. URL: https://doi.org/10.2991/978-94-6463-030-5_60.

7 Appendix

A1 Full Stock Analysis Table

*Verdict marked as "**EXCLUDED**" indicates that the asset was excluded from portfolio optimization due to insufficient model performance (ROC AUC < 0.5).*

Table 5: All stocks analyzed with XGBoost model

Stock	Verdict	P(>2%)	ROC AUC	PR AUC	Log Loss	Brier Score	Sector	Region
AAPL	BUY	0.5229	0.5551	0.6550	0.7019	0.2542	Technology	US
MSFT	HOLD	0.3524	0.6279	0.7738	0.7978	0.3007	Technology	US
NVDA	**EXCLUDED**	0.2646	0.3919	0.5324	1.4839	0.5054	Technology	US
TSM	SELL	0.2452	0.6257	0.8532	1.4081	0.4401	Technology	AS
SAP	**EXCLUDED**	0.4353	0.3093	0.6877	1.1638	0.4567	Technology	EU
AVGO	HOLD	0.4102	0.5915	0.7512	0.8009	0.3027	Technology	US
STM	BUY	0.5458	0.5695	0.3158	0.6415	0.2258	Technology	EU
ORCL	SELL	0.2892	0.5618	0.6961	0.8680	0.3309	Technology	US
INTC	BUY	0.5504	0.5626	0.5245	0.6933	0.2501	Technology	US
AMD	BUY	0.5520	0.5508	0.6612	0.8947	0.2952	Technology	US
BAC	BUY	0.6130	0.6822	0.8392	0.6125	0.2118	Finance	US
DB	**EXCLUDED**	0.2441	0.2696	0.6656	0.8947	0.3430	Finance	EU
HSBC	HOLD	0.4962	0.5919	0.7914	0.6923	0.2496	Finance	EU
JPM	HOLD	0.3407	0.5481	0.8629	0.8498	0.3252	Finance	US
MS	HOLD	0.4965	0.5563	0.6682	0.6881	0.2475	Finance	US
SAN	BUY	0.5207	0.6431	0.7950	0.7312	0.2630	Finance	EU
AXP	**EXCLUDED**	0.4538	0.4266	0.8029	0.7782	0.2916	Finance	US
BLK	SELL	0.1826	0.6265	0.7546	0.7551	0.2728	Finance	US
BNPQF	HOLD	0.4565	0.8015	0.8565	0.6455	0.2263	Finance	EU
GS	SELL	0.2490	0.6835	0.8708	0.9532	0.3620	Finance	US
ING	HOLD	0.4460	0.5087	0.6304	0.7106	0.2586	Finance	EU
LYG	**EXCLUDED**	0.0299	0.4854	0.6728	1.0693	0.3319	Finance	EU
UBS	HOLD	0.3366	0.5894	0.6561	0.7929	0.2969	Finance	EU
WFC	HOLD	0.3073	0.7443	0.8130	0.6712	0.2391	Finance	US
AMZN	BUY	0.5774	0.6564	0.8539	0.9438	0.3454	Consumer Staples	US
BUD	BUY	0.5306	0.6971	0.5770	0.6635	0.2352	Consumer Staples	EU
DEO	**EXCLUDED**	0.5016	0.4610	0.2958	0.9366	0.3518	Consumer Staples	EU
TM	HOLD	0.4383	0.6233	0.6552	0.6976	0.2471	Consumer Staples	AS
KO	HOLD	0.3354	0.5842	0.8060	0.6796	0.2432	Consumer Staples	US
KR	HOLD	0.3502	0.6976	0.7658	0.6137	0.2124	Consumer Staples	US
MDLZ	BUY	0.6150	0.6655	0.5454	0.6246	0.2165	Consumer Staples	US
NKE	BUY	0.8889	0.5454	0.4748	0.8160	0.2979	Consumer Staples	US
NSRGF	**EXCLUDED**	0.5706	0.3531	0.3046	0.7461	0.2761	Consumer Staples	EU

Stock	Verdict	P(>2%)	ROC AUC	PR AUC	Log Loss	Brier Score	Sector	Region
PEP	BUY	0.6101	0.6921	0.6670	0.6759	0.2384	Consumer Staples	US
PG	HOLD	0.4286	0.5946	0.6846	0.6879	0.2474	Consumer Staples	US
TSN	BUY	0.5708	0.7628	0.8038	0.6356	0.2214	Consumer Staples	US
UL	SELL	0.1354	0.8095	0.8542	0.6274	0.2202	Consumer Staples	EU
BP	**EXCLUDED**	0.4606	0.4924	0.4028	0.6928	0.2498	Energy	EU
COP	BUY	0.5394	0.5309	0.4032	0.7066	0.2567	Energy	US
CVX	**EXCLUDED**	0.6478	0.3880	0.4757	0.8215	0.3048	Energy	US
SHEL	BUY	0.5135	0.6407	0.6987	0.7078	0.2541	Energy	EU
SLB	SELL	0.2842	0.6289	0.4129	0.6418	0.2265	Energy	US
XOM	**EXCLUDED**	0.4365	0.3449	0.3832	0.7152	0.2608	Energy	US
DVN	BUY	0.6871	0.5215	0.4217	0.7256	0.2547	Energy	US
EOG	HOLD	0.4667	0.6910	0.6514	0.6713	0.2391	Energy	US
EQNR	**EXCLUDED**	0.4601	0.4968	0.4652	0.6959	0.2514	Energy	EU
HAL	**EXCLUDED**	0.3538	0.4351	0.2860	0.6405	0.2232	Energy	US
MPC	**EXCLUDED**	0.5124	0.4135	0.4520	0.6957	0.2513	Energy	US
OXY	HOLD	0.4317	0.5381	0.4292	0.6839	0.2454	Energy	US
VLO	**EXCLUDED**	0.2545	0.4711	0.4074	0.7644	0.2789	Energy	US
ABBNY	HOLD	0.3628	0.5210	0.6734	0.7921	0.2979	Industrial	EU
EADSF	HOLD	0.4899	0.5122	0.6655	0.7333	0.2699	Industrial	EU
GE	**EXCLUDED**	0.4756	0.3746	0.7634	0.7616	0.2835	Industrial	US
SBGSF	SELL	0.1124	0.6187	0.7148	1.3006	0.4718	Industrial	EU
TSLA	HOLD	0.3169	0.5723	0.5834	0.6883	0.2474	Industrial	US
CAT	**EXCLUDED**	0.4185	0.4855	0.5566	0.7080	0.2573	Industrial	US
ETN	**EXCLUDED**	0.3500	0.3715	0.7263	0.8226	0.3121	Industrial	EU
MMM	BUY	0.5606	0.5850	0.7618	0.6438	0.2236	Industrial	US
SDVKY	**EXCLUDED**	0.5500	0.3485	0.4225	0.7050	0.2559	Industrial	EU
META	HOLD	0.4317	0.6525	0.8834	0.7119	0.2592	Communication Services	US
CMCSA	**EXCLUDED**	0.4102	0.3260	0.3470	0.8598	0.3259	Communication Services	US
DYEGY	SELL	0.2208	0.5082	0.8132	1.1579	0.4501	Communication Services	EU
GOOGL	HOLD	0.4953	0.5297	0.7381	0.7151	0.2609	Communication Services	US
NFLX	SELL	0.2310	0.5510	0.8109	0.8117	0.3460	Communication Services	US
VOD	HOLD	0.4272	0.5611	0.8295	0.7293	0.2679	Communication Services	EU
DIS	HOLD	0.4687	0.7738	0.7186	0.6182	0.2134	Communication Services	US
ORANY	SELL	0.1983	0.7207	0.7478	0.6414	0.2208	Communication Services	EU
PSO	SELL	0.2888	0.6173	0.8159	0.9063	0.3474	Communication Services	EU
ITVPY	**EXCLUDED**	0.8886	0.4829	0.4783	1.0169	0.3369	Communication Services	EU
VZ	BUY	0.5846	0.6622	0.7811	0.6140	0.2126	Communication Services	US
T	**EXCLUDED**	0.5050	0.4140	0.8179	0.6625	0.2347	Communication Services	US
BAESF	**EXCLUDED**	0.3754	0.4757	0.6097	0.7990	0.2962	Aerospace & Defense	EU
RTX	HOLD	0.3603	0.6085	0.8773	0.6911	0.2490	Aerospace & Defense	US
AXON	**EXCLUDED**	0.2174	0.4530	0.7585	1.9472	0.6182	Aerospace & Defense	US
BA	BUY	0.5853	0.5794	0.4850	0.7351	0.2688	Aerospace & Defense	US
LHX	BUY	0.6238	0.6402	0.6802	0.6290	0.2192	Aerospace & Defense	US
NOC	BUY	0.6601	0.5624	0.6355	0.7009	0.2537	Aerospace & Defense	US
TXT	HOLD	0.3673	0.5206	0.5502	0.7531	0.2776	Aerospace & Defense	US
BEP	**EXCLUDED**	0.5334	0.3067	0.3193	0.7226	0.2647	ESG	US
ENPH	BUY	0.5566	0.6349	0.6765	0.7167	0.2617	ESG	US
FSLR	**EXCLUDED**	0.1338	0.4712	0.4620	1.0550	0.4620	ESG	US
IBDSF	**EXCLUDED**	0.4710	0.4130	0.5858	0.7054	0.2561	ESG	EU
PLUG	**EXCLUDED**	0.4773	0.4747	0.3455	0.7310	0.2688	ESG	US
VWSYF	HOLD	0.4604	0.5197	0.4791	0.7212	0.2624	ESG	EU
NEE	BUY	0.8328	0.5833	0.6536	0.7718	0.2739	ESG	US
SEDG	BUY	0.5225	0.5379	0.3734	0.7495	0.2778	ESG	US
RUN	**EXCLUDED**	0.5571	0.4354	0.3569	0.7140	0.2604	ESG	US
BLDP	HOLD	0.4608	0.5052	0.4909	0.7347	0.2688	ESG	US
DSDVF	HOLD	0.3847	0.7994	0.9094	0.6318	0.2196	Shipping	EU
FRO	**EXCLUDED**	0.4301	0.4043	0.4402	0.7045	0.2556	Shipping	EU
MATX	BUY	0.5718	0.6069	0.8011	0.6492	0.2282	Shipping	US
DAC	HOLD	0.4980	0.5779	0.6718	0.6876	0.2472	Shipping	EU
GOGL	BUY	0.6327	0.5897	0.7880	0.6297	0.2192	Shipping	US
KEX	**EXCLUDED**	0.4833	0.4600	0.5554	0.7021	0.2545	Shipping	US
SBLK	BUY	0.6537	0.6216	0.8054	0.6251	0.2174	Shipping	EU

A2 Portfolio Results

Portfolio Statistics – Markowitz Best Sharpe Ratio

Expected Return:	18.03%	Sharpe Ratio:	1.0298
Volatility:	17.51%	VaR (95%):	10.77%

Stock	**Weight**	**Sector**	**Region**
NOC	15.58%	Aerospace & Defense	US
AMD	13.57%	Technology	US
MDLZ	15.02%	Consumer Staples	US
DSDVF	19.41%	Shipping	EU
MS	4.65%	Finance	US
NEE	11.91%	ESG	US
MSFT	4.59%	Technology	US
KR	12.51%	Consumer Staples	US
SBLK	2.76%	Shipping	EU

Table 6: Asset allocation for the Markowitz Best Sharpe Ratio portfolio (red star in Figure 9).

Portfolio Statistics – Markowitz Minimum Volatility

Expected Return:	11.43%	Sharpe Ratio:	0.8049
Volatility:	14.20%	VaR (95%):	11.93%

Stock	**Weight**	**Sector**	**Region**
KO	8.34%	Consumer Staples	US
TM	13.9%	Consumer Staples	AS
DSDVF	9.76%	Shipping	EU
NOC	12.83%	Aerospace & Defense	US
SBLK	3.00%	Shipping	EU
RTX	9.82%	Aerospace & Defense	US
TSN	9.34%	Consumer Staples	US
KR	10.35%	Consumer Staples	US
BUD	11.35%	Consumer Staples	EU
PG	11.31%	Consumer Staples	US

Table 7: Asset allocation for the Markowitz Minimum Volatility (blue star in Figure 9).

Portfolio Statistics – Black-Litterman Minimum Volatility

Expected Return:	11.32%	Sharpe Ratio:	0.8172
Volatility:	13.85%	VaR (95%):	11.46%

Stock	Weight	Sector	Region
NOC	13.95%	Aerospace & Defense	US
VZ	12.44%	Communication Services	US
KR	23.77%	Consumer Staples	US
TM	18.4%	Consumer Staples	AS
SHEL	3.27%	Energy	EU
DSDVF	8.89%	Shipping	EU
PG	19.28%	Consumer Staples	US

Table 8: Asset allocation for the Black-Litterman Minimum Volatility Portfolio (blue star in Figure 10).

Portfolio Statistics – Black-Litterman Maximum Return

Expected Return:	29.19%	Sharpe Ratio:	0.6879
Volatility:	42.43%	VaR (95%):	40.61%

Stock	Weight	Sector	Region
GOGL	13.28%	Shipping	US
AMD	37.24%	Technology	US
TM	0.57%	Consumer Staples	AS
ENPH	36.57%	ESG	US
WFC	8.45%	Finance	US
SBLK	3.89%	Shipping	EU

Table 9: Asset allocation for the Black-Litterman Maximum Return Portfolio (green star in Figure 10).

A3 Code

The source code used in this thesis is available at GitHub.[1]

[1] https://github.com/smhovlan/Forecast-Informed-Portfolio-Optimization-A-Machine-Learning-Approach-to-Strategic-Asset-Allocation