Contents

1   General description

In a fast-evolving market environment, our strategy leverages advanced quantitative tools to identify return opportunities. We combine machine learning with economically motivated signals—U.S. macroeconomic calendar surprises, option-implied information (level, term, smile, volume, open interest), Financial Times trade/tariff news sentiment, and technical indicators on SPY and VIX—augmented by volatility-regime features (including rolled GARCH(1,1) parameters and persistence).

The objective is to improve the global performance relative to a buy-and-hold exposure to the S&P 500 through the SPY ETF. Daily predictions are translated into long positions in buckets of SPY's holdings. Our methodology's hedge is shown through transparent documentation of feature engineering, model selection, and evaluation.

2   Data and Feature Engineering

Macro Surprises In order to reflect the effect of macroeconomic surprises on the U.S. equity market, we built a set of features that capture the deviations between market expectations and realized economic data. These indicators quantify how new information entering the market diverges from what was previously anticipated, which we believe drives short-term movements in equity prices.

News Sentiment - Tariffs: We scrape around 10,000 articles from the *Financial Times* tagged to trade/tariffs from 2015 to 2025 and score sentiment with an LLM prompt constrained to output a 1–10 score (temperature 0 for determinism). We then aggregate the data to daily by averaging the sentiment scores of all articles in one day; missing days are forward-filled conservatively. We also define a direction measure that averages the direction of all articles in a day, bounded between 0 and 1. Finally, "maximum sentiment deviation" represents the article whose sentiment score diverges most significantly from the neutral value of 5, thereby giving an indicator of the day's most impactful content.

Options volatility term-structure: We incorporate a set of features derived from the US equity options market. Options data provide a forward-looking measure of investor expectations about volatility, risk, and sentiment. This is relevant for short-term predictions as option markets often react faster to new information.

Our set of options features capture three dimensions. First, the volatility term structure and smile. The slope of the implied volatility curve is measured both in the short-term and the longer-term. Similarly, skew is defined as the spread between OTM puts and OTM calls for both short- and long-term maturities. These measures reflect how investors price downside risk. Second, we estimate positioning and sentiment with put-call ratios based on open interest and volume. An elevated put activity relative to calls often reflects a higher demand for protection against downside risk, while a spike in call activity may indicate speculative risk-taking from investors. Finally, to account for changes in market positioning, we also include relative changes in OI and volume for OTM puts and calls.

GARCH features: We estimate a daily GARCH(1,1) model on SPY with a two-year rolling window and monthly refits, producing forward-looking annualized volatility, in-sample variance, and regime-persistence features $(\alpha_t, \beta_t, \alpha_t + \beta_t)$. These GARCH-based features are used as predictive inputs in the models to capture how returns behave across different volatility regimes, without directly affecting position sizing. In addition, we include (i) a volatility-regime dummy based on SPY/VIX (e.g., high-vol regime when rolling realized volatility exceeds a percentile threshold) and (ii) the rolled `rugarch` GARCH(1,1) parameters, one-step-ahead volatility forecasts, and persistence $(\alpha + \beta)$, capturing time-varying second moments of market risk.

Technical Indicators: Some technical indicators are calculated on the SPY : ATR, ADX, Aroon, Bollinger Bands, Chaikin Vol, CLV, EMV, MACD, MFI, SAR, SMI, Garman–Klass vol. We also compute a few technical indicators on the VIX: analogous momentum/level indicators. Indicators are lagged one step.

3   Building the strategy

Our trading strategy integrates multiple layers of predictive modelling, feature engineering, and portfolio construction to forecast U.S. equity markets' performance, which we capture using returns of the SPY ETF. The SPY's daily returns serve as the target variable, while a binary indicator of price direction is derived to enable both regression-based and classification-based prediction models.

Rather than investing in the SPY directly, we opted to expand our framework to include its constituents. We see this as a way to integrate higher levels of volatility into our strategy, and therefore capture higher upside potential. For each of the stocks, a rolling market beta is computed using a 180-day window, which we use to segment the constituents into 10 quantile-based buckets every day. This sorting allows us to adjust our exposure to individual stocks, depending on our prediction of SPY returns as a whole: we long higher-beta stocks when we forecast positive returns, and we long lower-beta stocks when our prediction is negative. This way, we capture a large amount of the upside when the overall SPY does well, while keeping a more defensive positioning otherwise.

Obtaining the predictions for our chosen set of models involves year-by-year rolling window testing from 2020 onward. For each year, the models are trained on the past 5 years of data, up to December 31 of the preceding year. Both regression and classification models are considered, encompassing a broad spectrum of statistical and machine learning techniques: Ordinary Least Squares (OLS), Logistic regression, Stepwise regression, LASSO, Relaxed LASSO, Ridge regression, Bagging and Random Forest (regression and classification variants).

For each of the tested models, we aggregate their predictions weekly over the complete testing period. From there, we build the following two strategies:

- *Max Sharpe*: a portfolio consisting of all models that outperform the SPY, based on their respective Sharpe ratios as a risk-adjusted performance measure.

- *Weighted Alpha*: a performance-oriented portfolio, which is composed of the models with higher average weekly returns than the benchmark.

Each composite strategy is computed as a weighted combination of individual model returns, where the weights are computed as follows:

$$\sum_{\substack{j \\ \mathsf{X}_j > \mathsf{X}_{\mathsf{SPY}}}} \frac{w_j}{w_N}, \quad \text{where } N = \#\{j : \mathsf{X}_j > \mathsf{X}_{\mathsf{SPY}}\}, \quad \text{and } \mathsf{X} \in \{\text{Sharpe ratio}, \text{average returns}\}.$$

For each of the strategies, we visualize their cumulative growth trajectories and compare them to our baseline buy-and-hold portfolio.

4   Recommendation and limitations

After comparing several approaches, we observe that the results of the Max Sharpe and Weighted Alpha portfolios drastically outperformed the SPY in the out-of-sample period. Machine learning models such as Random Forest and bagging performed particularly well, followed by penalized regression models (Ridge,

Lasso, Relaxed-Lasso), and finally OLS, Logistic regression and Stepwise regression model.
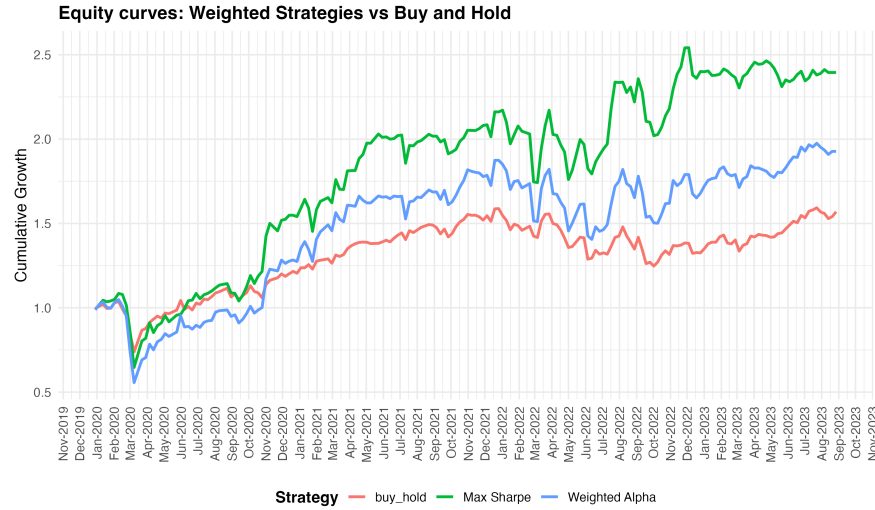

Figure 1: Cumulative Performance

Model	Annual Return	Annual Vol	Annual Sharpe
OLS	5.6%	36.1%	0.153
Bagging	34.4%	34.5%	0.913
LASSO	25.1%	42.4%	0.552
Logit (class.)	20.8%	38.8%	0.496
Random Forest (reg.)	39.0%	35.4%	0.884
Relaxed LASSO	25.1%	42.4%	0.552
RF (class.)	22.4%	39.3%	0.531
Ridge	26.7%	42.4%	0.582
Stepwise AIC	14.2%	37.5%	0.355
Buy & Hold SPY	16.4%	21.7%	0.695
Max Sharpe	48.9%	51.1%	0.824
Weighted Alpha	30.5%	45.1%	0.610

Table 1: Annualized performance (2020–01 to 2023–09).

These strategies can be adapted to different risk profiles. Risk-averse investors may prefer the SPY ETF, as none of the machine learning strategies achieved lower variance due to the nature of our beta strategy. On the other hand, risk-seeking investors aiming to generate alpha over SPY might prefer the Max Sharpe portfolio, which allocates capital to strategies with higher Sharpe ratios. However, given its limited diversification (only two models in the portfolio), the Weighted Alpha strategy offers a more balanced alternative with broader model exposure.

It is important to note that while the S&P500 is generally considered as diversified, our beta-bracket strategy may not provide sufficient diversification across sectors due to the fact that high-beta and low-beta brackets have a smaller number of stocks and are not exceedingly dynamic on each rebalancing period. Finally, some input features such as macroeconomic surprises or news sentiment are naturally noisy and subject to the information provider or timing biases, which can affect signal reliability over time.