1. Project Overview

The stock market is notoriously hard to forecast due to its reflexivity between the actions of market participants and their perceptions. Strategies such as active value investing which have outperformed in the past have struggled to maintain their performance for well over a decade now as passive investing takes the lead. Momentum and quantitative trading strategies which worked under certain market regimes can find themselves struggling with capacity constraints, timing issues, and subject to nasty volatility skews, fat kurtosis tails, and huge margin calls. What works in one stock market may not work in another and what works now may not work in the future. Consequently, it has been empirically found that more than 90% of traders lose money in the long run and more than 90% of active managers do not outperform their benchmarks (the S&P500 index being chief among all) once all costs have been factored in.

2. Problem Statement

All market participants want to make money. Even those who are in the market for hedging risk want to be overall profitable. To be profitable, one needs to be holding the right position (long or short) for the future state of the market. The problem is that the stock market cannot be easily forecast. The further the horizon, the more inaccurate forecasts nearly always are. Short-term forecasts can however be made with a reasonable amount of accuracy if the right tools are used. This project seeks to answer the question: Is it more profitable for a trading/investing strategy to predict the current market regime (bear or bull market) or predict the forward price returns in advance?

3. Research Methodology

Asset	SPY ETF (ticker: "SPY")
Underlying index	S&P500 Index
Model Training Period (in-sample)	Since Inception 1993-2009 (except in the case of Hidden Markov Regime Prediction, three different training period approaches were used: 1) Static-Window between 1993-2009, 2) Expanding-Window that accumulates each passed year as part of its training period, 3) Walk-Forward Window that rolls into a new 5-year training period at the end of every year)
Model Testing Period (out-of-sample)	2010-2023/1 or shorter depending on the duration of timeseries features being used
Model Metric	Cumulative returns in % during the backtesting period that yields the maximum positive returns

“Best Model” Approach

Both supervised and unsupervised machine learning(“ML”) tools will be used to try and predict for both the current market regime and forward price returns. Supervised ML approaches used include LightGBM and ARIMA modeling while unsupervised ML uses only Hidden Markov modeling tools.

Once the “best model” has been found for each of the two tasks, the model’s forecasts (“trading signals”) will be backtested out-of-sample to measure each model’s profitability during the backtesting period between 2018-2023/1.

The approach that yields the highest cumulative returns during the backtesting period will be concluded as the preferred strategy to generate alpha returns.

Feature Selection

There are two types of features used to train the models. Depending on the model and context in which it is being used, the feature set may not be the same.

1) Technical Analysis Indicator Features generated by ‘*ta-lib*’ and ‘*ta*’ python packages:
   A. Trend
      a. Mass Index
      b. Average Directional Index (“ADX”)
   B. Momentum
      a. Relative Strength Index (“RSI”)
      b. Donchian Channel
   C. Return
      a. 1-Month Returns
      b. 3-Month Returns

2) Timeseries Features generated using either ‘*tsfresh*’ package or organically augmented on a rolling-window approach

Backtesting Approaches

1) Market Regime Prediction Backtest Approach

Each trading day, we look back at the market regime signal generated at the end of the last trading day to decide whether to buy/sell SPY ETF today. If the last regime signals that the current market state is bullish, we go long the asset; if it signals bearish, we go short the asset. The trading strategy is a reactive one rather than anticipatory.

2) Predicted Forward Returns Backtest Approach

This is a day trading strategy where all open positions will always be closed at the end of the day. At the start of each trading day, once the OPEN price is available, the model will use that to predict the CLOSE price for that day. If the CLOSE is higher than OPEN, assume buying at the OPEN price and selling at the CLOSE price. If CLOSE is lower than OPEN, assume selling at the OPEN price and buying at the CLOSE price.

4. Research Findings

Part A: Market Regime Prediction

A.1: Hidden Markov Model(“HMM”) Clustering

Among the three approaches of model fitting, the Static-Window training model has been found to produce the most stable market regime classification (“hidden states”) across the testing period. While each approach has its strengths and weaknesses, the Static-Window trained model also produces the most profitable backtest results (Fig 2.). For this reason and the reliable persistence of HMM hidden state classification, we will use the market regime labels predicted by this model as a feature input for LightGBM and ARIMA Forecasting to check if the autoregressive components of such a timeseries can be used as a forecasting input (Fig 1.).

Fig 1. Hidden States Predictions during Test Period (where 0=bull regime; 1=bear regime)

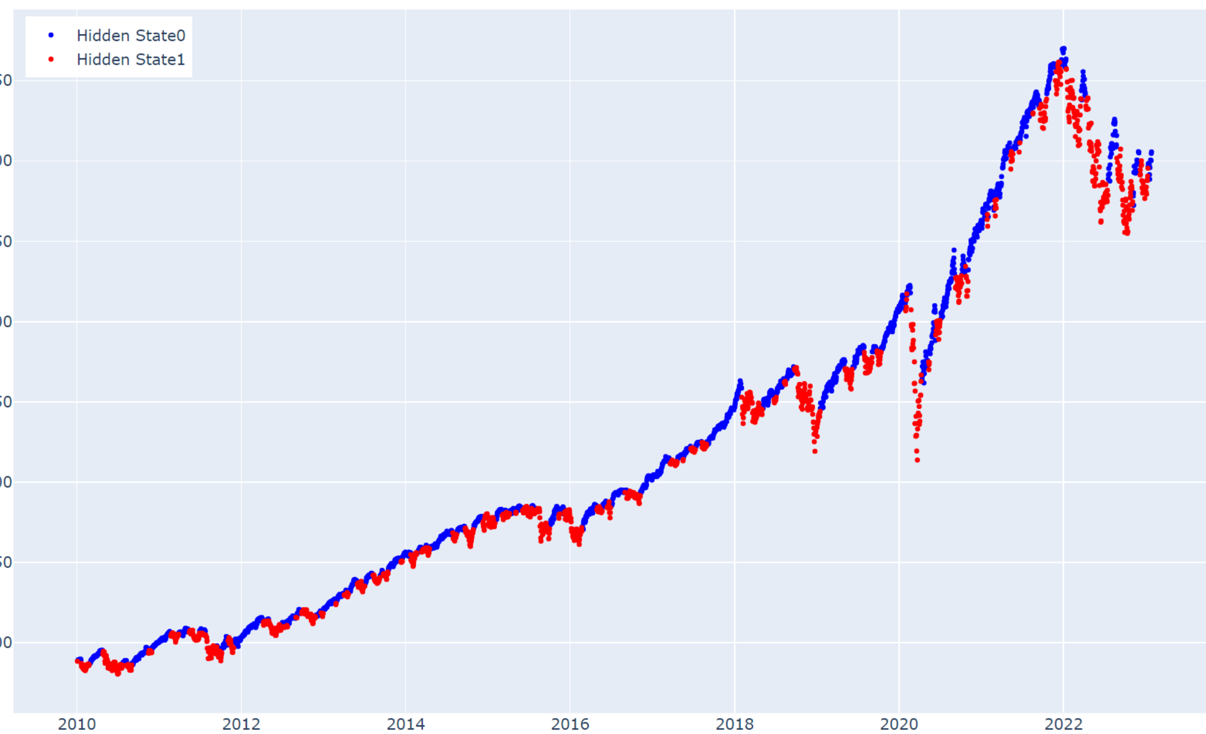


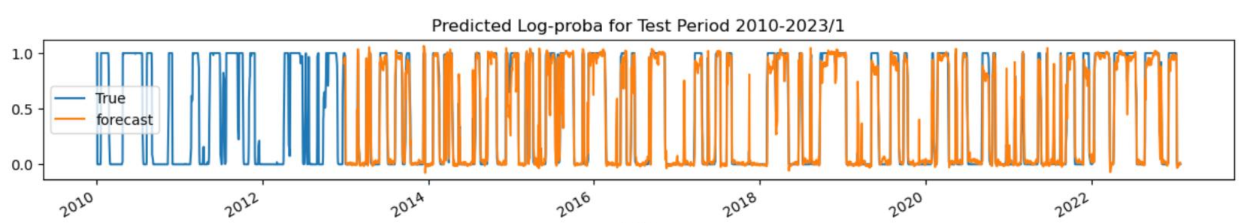
Fig 2. Backtest Profitability between 2018-2023/1

A.2: LightGBM Forecasting

Two different approaches have been tried here.

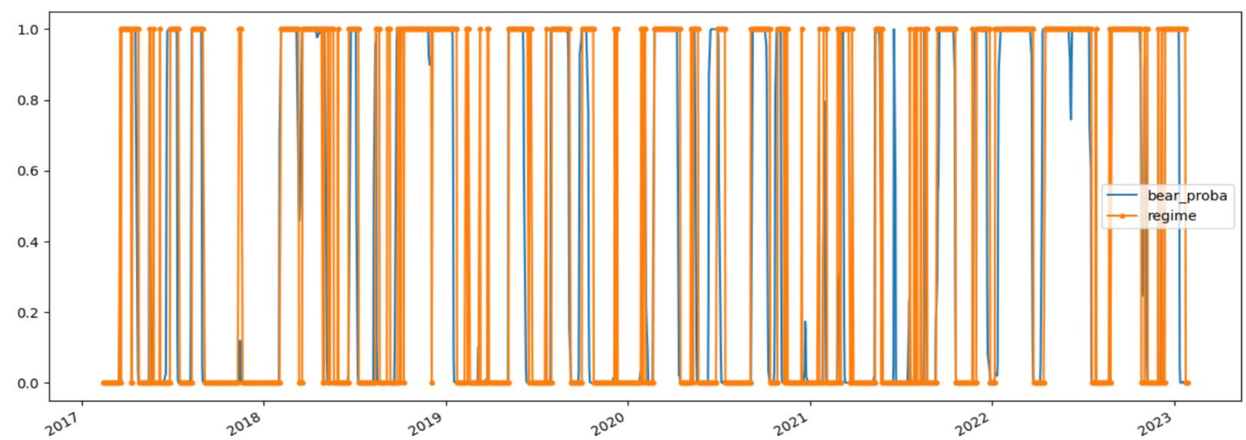
First, using the log-probabilities generated by the previous HMM Static-Window trained model as the feature input, the LGBM model has been used to forecast one-step ahead and multi-steps ahead. The result is that the LGBM's one-step ahead forecasts mostly match the HMM's regimes (Fig 3.) whereas the multi-steps prediction is completely off-the-mark in the test period.

Fig 3. One-Step Ahead Predicted Log Probabilities during Test Period (first 750 days used as inputs)



Second, instead of using HMM-generated log probabilities for regime classification, timeseries features derived from just 1-Month Returns were used as the input to train and predict the HMM's regimes. Similarly, the regime predictions closely matched the HMM's regimes in the test period (Fig 4.).

Fig 4. One-Step Ahead Predicted Log Probabilities during Test Period



While visually it may seem that both approaches were generally consistent with the HMM's regime classification, the actual backtest results from this LGBM forecasting strategy turned out to be negative due to quite a significant number of false positives/false negatives causing substantial drag on the trading performance. (Fig 5 and Fig 6)

Fig 5. Backtest Profitability for Fig 3 predictions between 2018-2023/1

Fig 6. Backtest Profitability for Fig 4 predictions between 2018-2023/1

A.3: ARIMA Forecasting

The use of ARIMA and its variant SARIMA was found to be largely a futile exercise. While it produces very accurate forecasting in-sample during the training period (Fig 7.), it failed to yield any meaningful results out-of-sample during the test period (Fig 8.). For this reason, no backtesting was done for this strategy.

Fig 7. Basic ARIMA Regime State Predictions during Training Period (showing only first 500 days of the period)

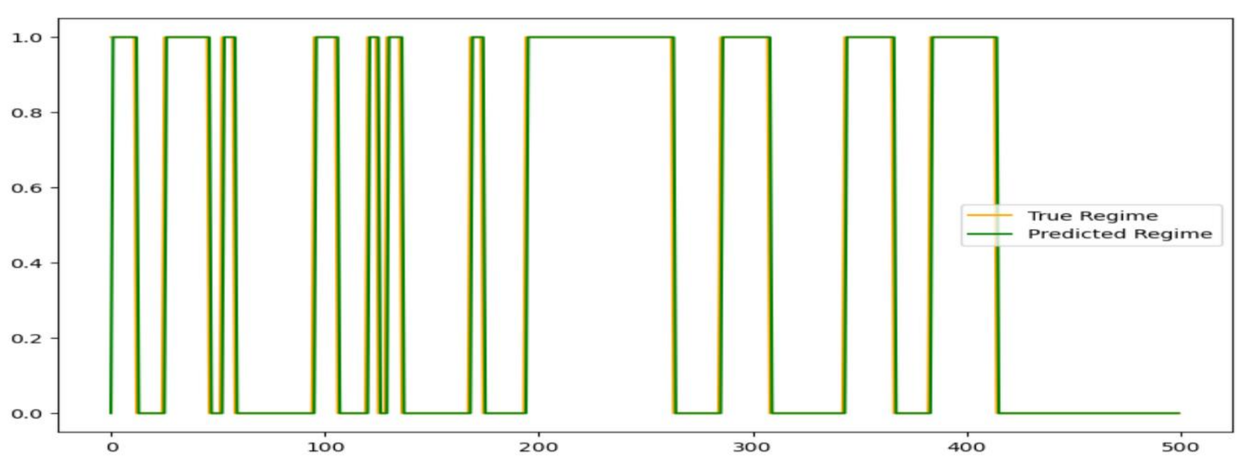
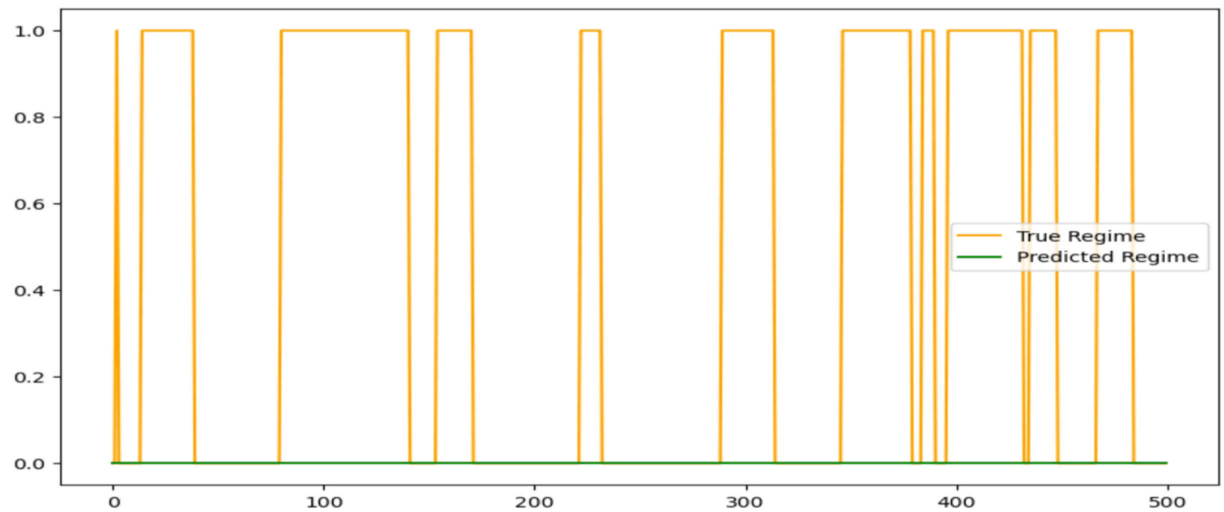


Fig 8. Basic ARIMA Regime State Predictions during Test Period (showing only first 500 days of the period)



Part B: Forward Price Returns Prediction

B.1: LightGBM Weekly Returns Forecasting

The use of either highly relevant technical indicators or timeseries features derived from just 1-week returns failed to yield reasonably accurate forward 1-week returns forecast (Fig 10 and Fig 11.). While both models were generally able to get the volatility of returns aligned with actual 1-week realized volatility, the directional forecasts were largely off. Consequently, no backtesting was carried out for LGBM forecasting since it is plainly obvious that the predicted forward returns failed to conform with the actual realized returns.

Fig 10. Predicted 1-Week Returns vs Realized 1-Week Returns using TA Features during Test Period

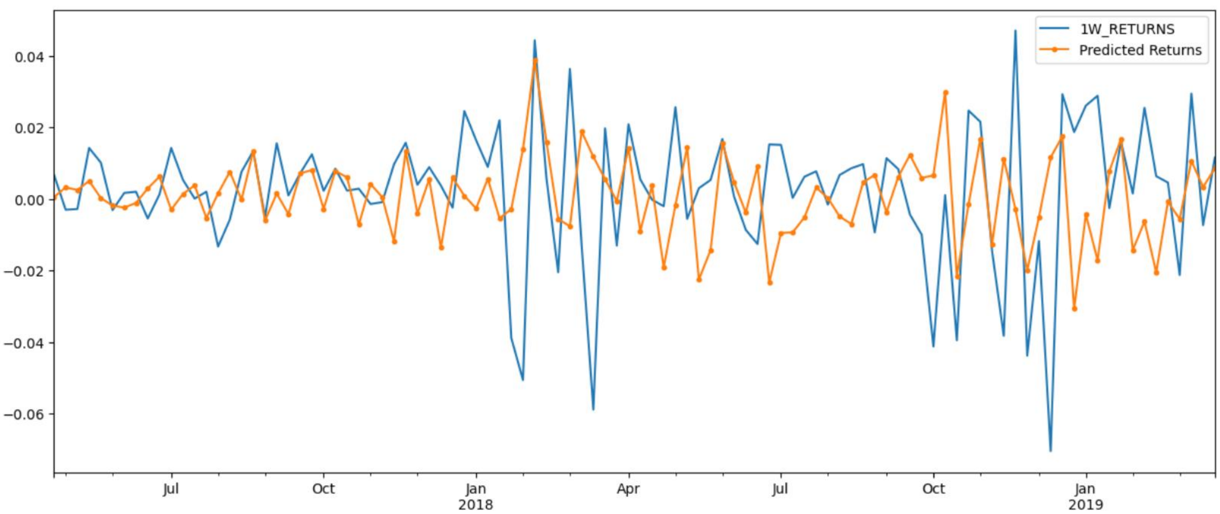
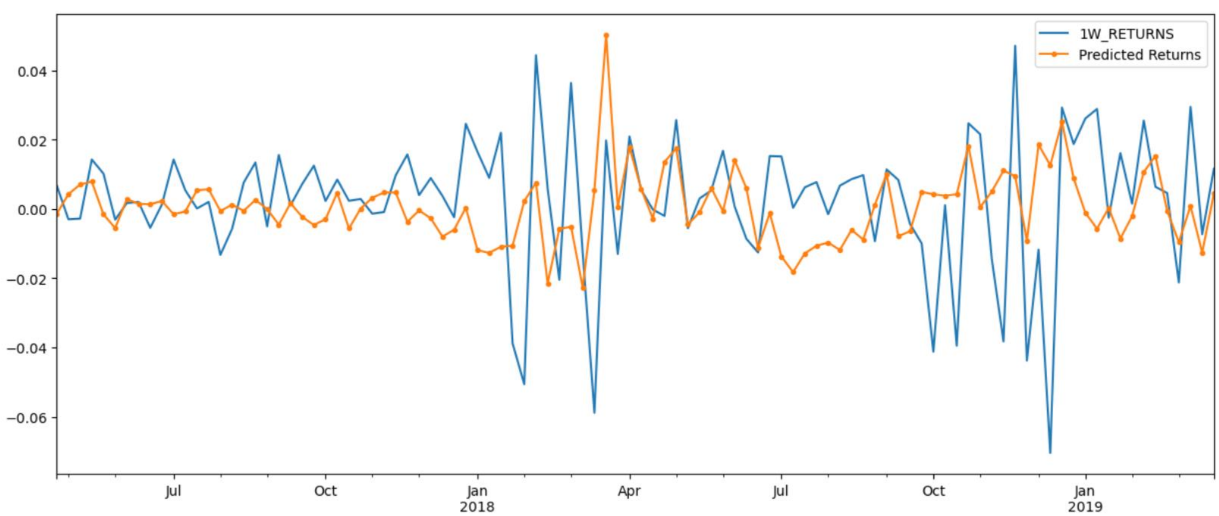


Fig 11. Predicted 1-Week Returns vs Realized 1-Week Returns using Timeseries Features during Test Period



B.2: Hidden Markov Model Daily Returns Forecasting

The number of hidden states found to yield reasonably accurate forward Friday's Close prices is 10 (Fig 12.). However, it appears that the predicted CLOSE prices were mostly off-the-mark and do not match the actual CLOSE accurately therefore, resulting in very bad trading performance during the backtest period (Fig 13.). As the backtest assumed one is able to trade at the exact OPEN and CLOSE prices to derive the results, the reality of trading execution renders this impractical and therefore, all slippage accounted for, the actual backtest results will be much more negative than shown in Fig 13.

Fig 12. Predicted Close Price vs Actual Close Price during Backtest Period (Zoomed-in for 2018-2019)

Fig 13. Backtest Profitability between 2018-2023/1

5. Research Limitations

1. Trading Costs such as short-financing costs, trading commissions and liquidity costs have not been simulated in the backtesting.
2. Use of non-timeseries feature inputs may not be sufficient to capture underlying statistical patterns
3. Models such as the HMM-clustering of market regimes did not work as well for SPY ETF compared to other traditionally more volatile assets such as oil and gold. Even the choice of training window (static window, expanding window or walk-forward window) can result in vastly different backtesting profitability depending on the asset. To obtain a statistically sound profitable trading strategy, one needs to not only select the right model but also the right training window.
4. Due to the inconsistencies of quantitative strategies, one needs to be especially clear about the characteristics of an asset that allowed the strategy to work and re-train the model's use if such characteristics changes in the future.

6. Conclusion and Recommendations

Hidden Markov models proved to be reasonably capable of predicting current market regimes but underperformed considerably when trying to predict forward price returns.

For the best alpha strategy, it is recommended to employ HMM models to predict the current market regime and not trade unnecessarily.