SCHOOL OF OPERATIONS RESEARCH AND INFORMATION ENGINEERING
COLLEGE OF ENGINEERING
CORNELL FINANCIAL ENGINEERING, MANHATTAN
CORNELL UNIVERSITY

Financial Engineering Project

REGIME DETECTION FOR PORTFOLIO OPTIMIZATION

Presented to the Faculty of the Graduate School of Cornell University
in partial fulfillment of the requirements for the Master of Engineering Degree
and the Financial Engineering Concentration

By:
Enming Zhang (ez228)
Haobing Meng (hm458)
Huimin Wang (hw447)
Joel Dsouza (jnd74)
Niharika Dalsania (nd322)
Yiran Hao (yh494)
Yuanping Du (yd346)

Faculty Advisor:
Prof. Sasha Stoikov

December 2022

Faculty Advisor's Signature: Date: 12.12.2022

ABSTRACT

Financial markets are dynamic and unpredictable. Understanding, recognizing and predicting trends are crucial because of their connections to economies, asset valuations, political turmoils and fiscal policies. While the statistical approaches have been the main approach of regime switching models, there have been attempts to detect regimes using machine learning methodologies. In this paper, we use unsupervised learning in the form of Gaussian Mixture Model (GMM) to identify clusters/regimes in the financial market using macro factors since 2002 and test the model from 2015 to 2020. The regimes detected show persistence in their properties throughout history and therefore are labeled accordingly. We demonstrate our framework's effectiveness by evaluating the performance of trading strategies with and without regime detection and conclude with our results that regime detection helps boost overall performance by leveraging the fact that certain assets do well in certain market conditions.

1. BACKGROUND

Financial markets are ever-changing and predicting returns of assets is a difficult task. Wise investment decisions therefore require a great judgment on the current market conditions and their alterations. While certain alterations/regimes may only last a short while, others often last for extended periods. Modeling various market regimes can be an effective tool, as it can enable macroeconomically aware investment decision-making in market conditions like Bull, Bear, high inflation, etc. Furthermore, risk management must take into account how risk premiums, volatility, and correlation change over time in these regimes. The recent unrest brought on by the Covid financial crisis serves as an example of the effects of a paradigm shift in investors' perceptions of risky assets. It therefore brings us to a fundamental question - What regime are we in?

A data-driven approach to regime detection involves using historical data on assets to define the regimes. A particular illustration of this approach is the Gaussian Mixture Model (GMM) for clustering and then labeling the clusters based on the asset's mean returns and standard deviations as discussed in Two Sigma's Machine learning approach to regime detection. The study presented in our paper will build on Two Sigma's regime detection approach by also detecting out-of-sample regimes, building optimal portfolios of each regime, backtesting trading strategies and changing model settings to compare performance.

The study is divided into different sections. First, we start with introducing the factors taken for our study and the rationale behind it. We then talk about the data preprocessing and initial analysis to construct features for the model. We briefly explain the Gaussian Mixture Model and the hyperparameters involved. Further, the entire research pipeline is explained and each component is described with the results. Also, we examine if regime detection adds value to building portfolios by comparing their results of the performance using several metrics. We conclude with our recommendations/takeaways along with discussing future scope of the study.

2. DATA

The following subsections discuss the details of the factor data used in this study and the steps taken to preprocess all the data. All the data used here ranges from January 2000 to August 2022.

2.1. FACTOR INTRODUCTION

A total of 15 factors are used here, which can be broadly divided into four big categories: Core Macro, Secondary Macro, Macro Styles, and Equity Styles. These factors are introduced in Table 1 below.

	Factor	Description
Core Macro	Equity	Exposure to the long-term economic growth and profitability of companies
	Interest Rates	Exposure to time value of money
	Credit	Exposure to corporate default and failure-to-pay risks specific
	Commodities	Exposure to changes in prices for hard assets
Secondary Macro	Emerging Markets	Exposure to the sovereign and economic risks of emerging markets relative to developed markets
	Foreign Currency	Exposure to moves in foreign currency values versus the portfolio's local currency
	Local Inflation	Exposure to inflation-linked rates relative to fixed nominal rates within the local currency area
	Equity Short Volatility	Negative exposure to the moves in equity market volatility
Macro Styles	Fixed Income Carry	Exposure to high-yielding 10-year bond futures funded by low-yielding 10-year bond futures
	Foreign Exchange Carry	Exposure to high-yielding G10 currencies funded by low-yielding G10 currencies

	Factor	Description
	Trend Following	Long-short exposure to multi-asset-class futures based on 6- to 12-month trailing returns
Equity Styles	Low Risk	Exposure to stocks with low market betas and residual volatility funded by higher-risk stocks
	Momentum	Exposure to stocks that have outperformed recently funded by underperforming stocks
	Quality	Exposure to stocks with low leverage, stable and high-quality earnings, and high profitability and investment quality, funded by lower-quality stocks
	Value	Exposure to stocks with low prices relative to accounting fundamentals and past prices, funded by higher-priced stocks

Table 1: Factor Introduction

2.2. FACTOR PREPARATION

In following subsections we present the details of factor construction along with necessary preprocessing and factor analysis to get the final dataset for the study.

2.2.1. DATA SOURCE

The raw price series for all the 15 factors introduced earlier is fetched from Bloomberg using a combination of tickers, details of which are discussed next.

2.2.2. FACTOR CONSTRUCTION

The following Table 2 shows the tickers used for fetching raw price series corresponding to all factors and details of how these price series are combined to get final factor data.

	Factor	Construction Details	Bloomberg Tickers
Core Macro	Equity	MSCI All Country World Index	MXWD
	Interest Rates	Bloomberg Barclays Global Government 7 to 10 Years Hedged to USD	LGY7TRUH
	Credit	Average of four indices: Bloomberg Barclays Pan-European High Yield (EU HY), Bloomberg Barclays Euro Aggregate Corporate (EU IG), Bloomberg Barclays US Corporate High Yield (US HY), and Bloomberg Barclays US Corporate (US IG)	LP01TREU LECPTREU LF98TRUU LUACTRUU
	Commodities	S&P GSCI index, a benchmark for the commodity markets	SPGSCI
Secondary Macro	Emerging Markets	Average of two indices: MSCI Emerging markets Index (EM equity), and Bloomberg Barclays Emerging Markets USD Aggregate (EM Credit)	MXEF EMUSTRUU
	Foreign Currency	Calculating returns to holding a diversified basket of G10 currencies relative to USD	AUD, CAD, CHF, EURO, GBP, JPY, RMB
	Local Inflation	Average of two indices: Bloomberg Barclays US Treasury Inflation-Linked Notes, and Bloomberg Barclays Euro Government Inflation-Linked Bond All Maturities	LBUTTRUU BEIGIT
	Equity Short Volatility	CBOE S&P 500 PutWrite Index	PUT

	Factor	Construction Details	Bloomberg Tickers
Macro Styles	Fixed Income Carry	DB Rates Diversified Strategy 016 USD Index	DBDRC3U3
	Foreign Exchange Carry	DB CoreSeries FX Carry Balanced USD Index	DBFXCAR1
	Trend Following	DB Cross Asset CTA Trend Index	DBCAUCTA
Equity Styles	Low Risk	DB Equity Low Beta Factor 2.0 USD Excess Return Index	DBRPGEBU
	Momentum	DB Equity Momentum Factor 2.0 USD Excess Return Index	DBRPGENU
	Quality	DB Equity Quality Factor 2.0 USD Excess Return Index	DBRPGEQU
	Value	DB Equity Value Factor 2.0 USD Excess Return Index	DBRPGEVU

Table 2: Factor Construction

2.2.3. DATA PREPROCESSING

The data series for all factors ranged from January 2000 and to August 2022. We align the dates of all the factors and fill the missing data using forward-filling. Then, we calculate the log returns of the price series to get a stationary time series, a very important preprocessing for financial data. We average the required returns series for factor construction. For instance, for the credit factor, we first calculate the log returns of EU IG, EU HY, US IG, and US HY respectively, and then compute the credit factor using the average of the returns of these for indices. Similarly for other factors that are constructed using the average of multiple returns series. Then, we perform a correlation analysis and ensure that the factors used are fairly uncorrelated. Figure 1 shows the correlation heatmap.

The correlations between factors are not very significant since the majority of the plot is deep color, which indicates a correlation close to 0. Some pairs of factors have high correlation. For instance, Equity and Equity Short Volatility have a correlation of 80% and Equity and Emerging Market have a correlation of 73%. These high correlations mean that one asset return might depend on another asset return. But these high correlations are rare and we can broadly conclude that the factors are orthogonal, hence making a good base for the factor data. However, we further conduct principal component analysis to verify our conclusion.

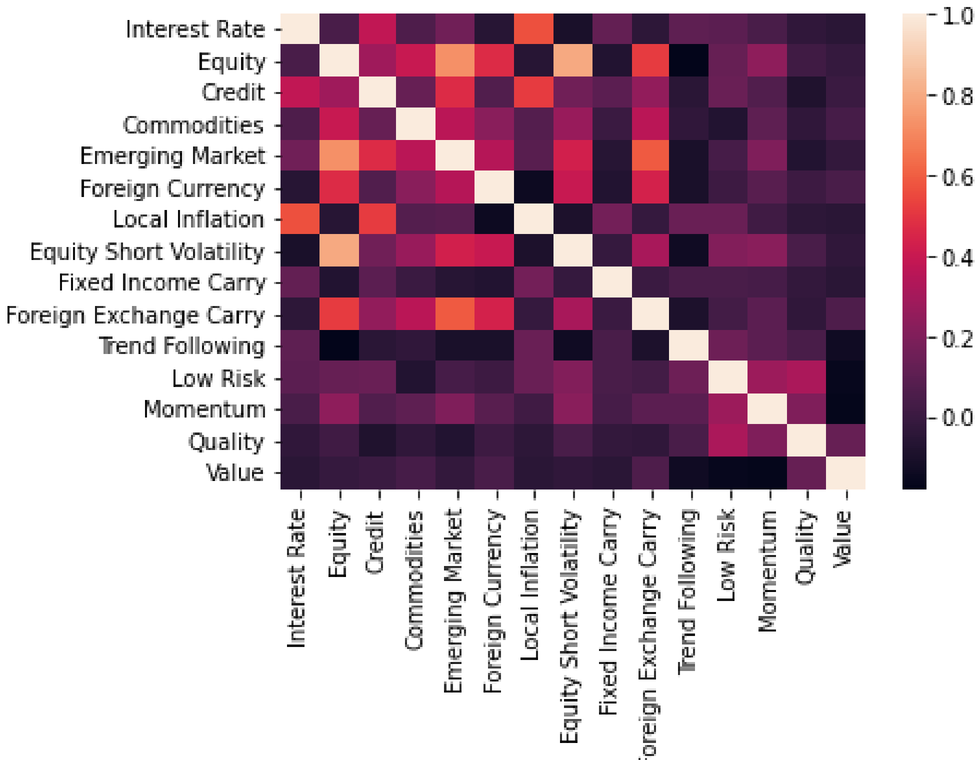

Figure 1: Correlation Analysis

2.2.4. PRINCIPAL COMPONENT ANALYSIS

To check the orthogonality of the factors, we conduct a Principal Component Analysis (PCA) on the 15 factors. Figure 2 is the cumulative explained variances of the first 12 principal components. We can see that in order to explain 80% of the variance, we have to use the first 8 principal components. In addition, to explain 90% of the variance, we need the first 12 principal components. Almost every factor contributes to the cumulative variance. It indicates that these factors are orthogonal and every factor contributes some unique information to the model. So there is no need to eliminate any factor or use principal components as factors in our study.

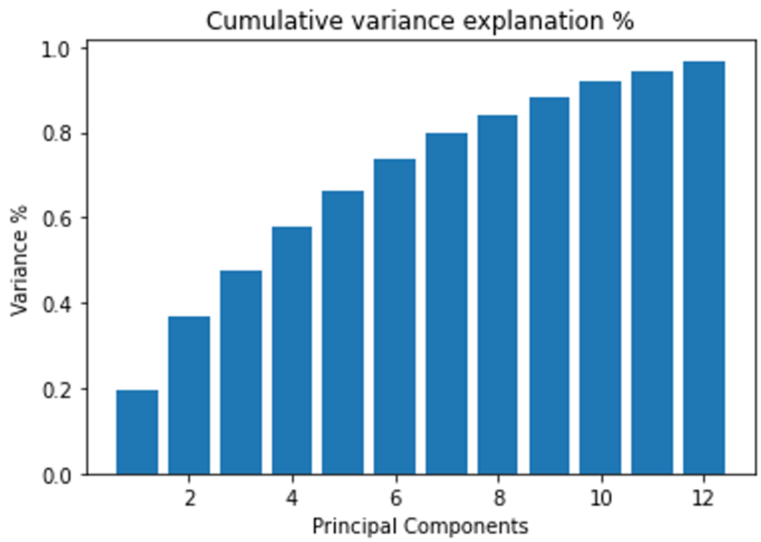

Figure 2: Variance Explanation by Top Principal Components

3. RESEARCH PIPELINE

In this study, we not only look at regime detection and interpretation, but also develop regime based portfolios and investment strategies to boost overall performance. Figure 3 shows a flow chart that summarizes the structure of the entire research pipeline.

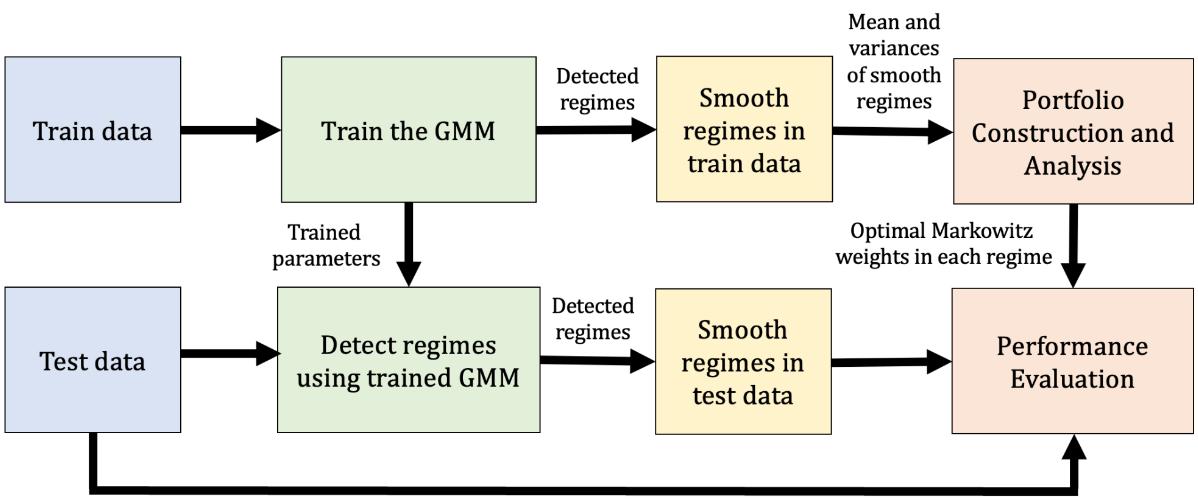

Figure 3: Research Pipeline

Following Figure 4 shows the details of each component shown in our research pipeline. The implementations and analysis for each will be discussed in the following sections.

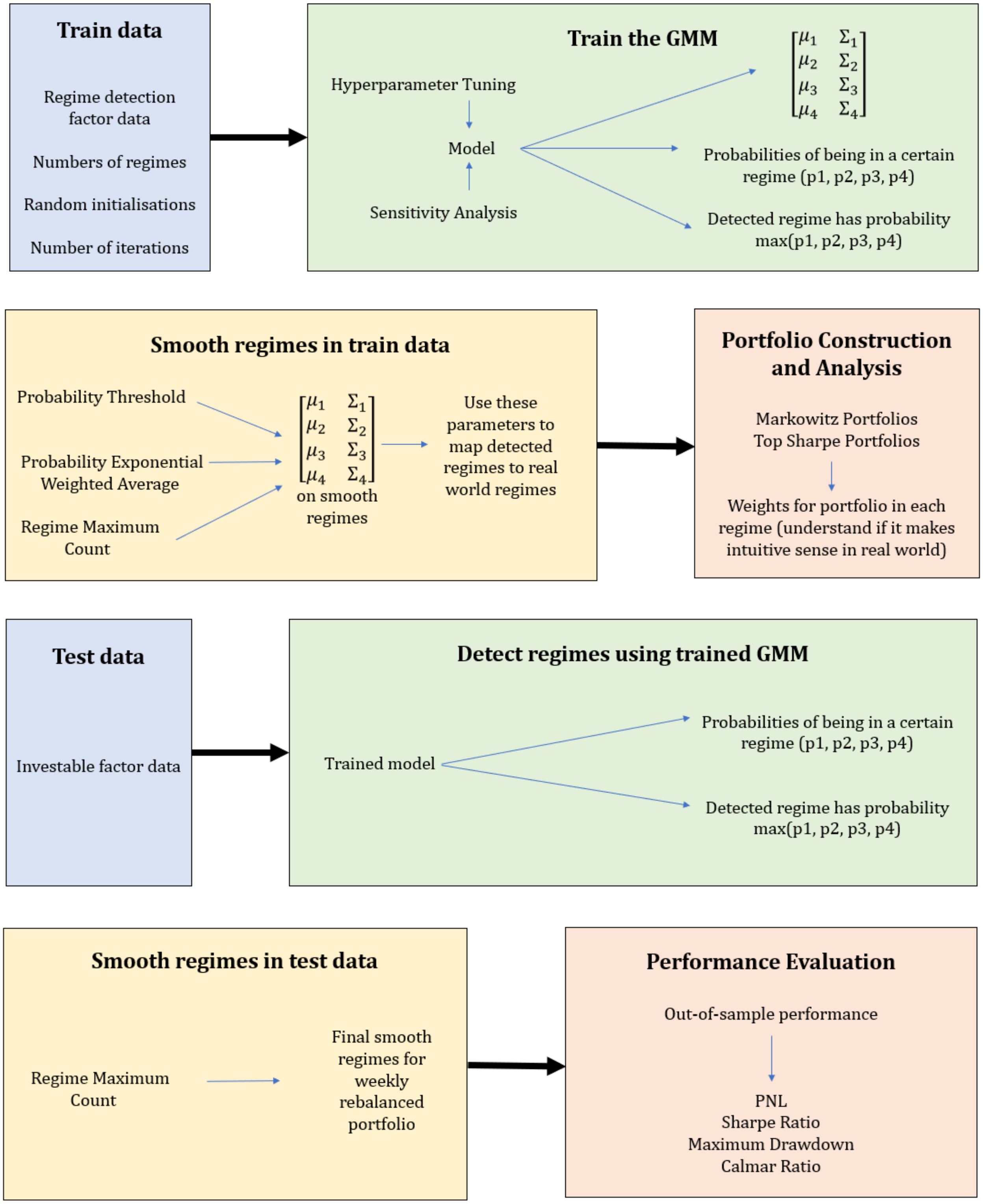

Figure 4: Details of Research Pipeline

Initially, we use the training factor data to fit a GMM model that gives the daily regimes. However, the regime changes very frequently which would cause potential issues, hence arising the need for regime smoothing. Then, we collect the regime-wise factor returns on the days that have been classified as a particular regime based on the smoothed regimes, and calculate the regime-wise mean and covariance matrix. First, we make sense of these regimes and map them to real world plausible regimes based on their return characteristics. Then, we build regime-wise Markowitz and Top Sharpe portfolios for our investment strategies. Then, we apply the trained GMM model to our test data to classify the regimes in the test period, and smooth the test regimes, similar to training period. Then for each week, we hold the portfolios built earlier, corresponding to the current regime. We apply multiple metrics to evaluate the performance of our regime-switching strategies.

4. REGIME DETECTION MODEL

In the following subsections we look at the working of the regime detection model and make sense of the identified regimes in terms of mapping them to real world scenarios.

4.1. GAUSSIAN MIXTURE MODEL

Gaussian Mixture Model (GMM) is a probabilistic unsupervised model that assumes all the observations are generated from a mixture of a finite number of Gaussian distributions with unknown parameters, shown in Figure 5. It is like k-means clustering, but including information about centers and covariance of the data and latent Gaussians.

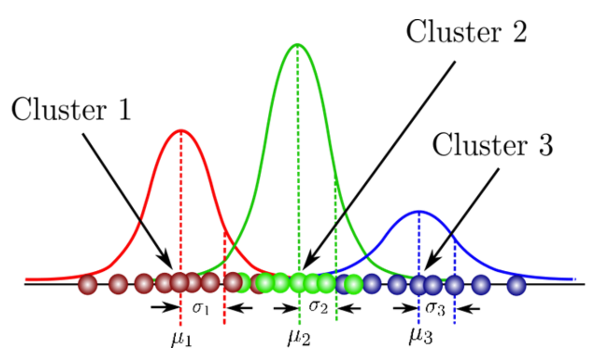

Figure 5: Illustration of Gaussian Mixture Model

Suppose there are in total K Gaussian mixtures for the GMM. Then GMM is parameterized by the distribution of each gaussian component and the probability that each observation belongs to each gaussian component. Given N observations in total,, the aim of GMM is to solve the parameters, discussed in Table 3, that maximizes the log-likelihood given the observations.

Parameter	Mathematical Expression
The distribution of each gaussian component - mean and covariance	$\theta = \{\theta_1, \theta_2, ..., \theta_K\}$ $\theta_k = \{\mu_k, \Sigma_k\},\ k = 1, 2, ..., K$
The probability that each observation belongs to each gaussian component	γ_{jk} (the probability that the j-th observation belongs to the k-th Gaussian)
The probability that all observations come from the k-th Gaussian	$\alpha_k = \sum_{j=1}^{N} \gamma_{jk}$
The probability density function for the k-th Guassian component	$\phi(x \mid \theta_k) = \frac{1}{(2\pi)^{D/2}\lvert\Sigma_k\rvert^{1/2}} exp(-\frac{(x-\mu_k)^T\Sigma^{-1}(x-\mu_k)}{2})$
The mixed probability density considering all K components	$P(x \mid \theta) = \sum_{k=1}^{K} \alpha_k \phi(x \mid \theta_k)$
The likelihood given the Gaussian parameter	$L(\theta) = \prod_{j=1}^{N} P(x_j \mid \theta),\ x_j$ denoting observation j
The objective function of GMM	$max\ log L(\theta) = \sum_{j=1}^{N} log\ P(x_j \mid \theta)$

Table 3: Defining parameters of Gaussian Mixture Model

GMM estimates parameters using Expectation–maximization (EM) as follows:

- Initialize the params
- E-Step: Given the current parameters, calculate the probability that the j-th observation comes from k-th model:
    - $\gamma_{jk} = \frac{\alpha_k \phi(x_j \mid \theta_k)}{\sum_{i=1}^{K} \alpha_i \phi(x_j \mid \theta_i)};\ j = 1, 2, ..., N;\ k = 1, 2, ..., K$
- M-Step: Calculate the updated params:
    - $\mu_k = \frac{\sum_{j=1}^{N} (\gamma_{jk} x_j)}{\sum_{j=1}^{N} \gamma_{jk}},\ \Sigma_k = \frac{\sum_{j=1}^{N} \gamma_{jk}(x_j-\mu_k)(x_j-\mu_k)^T}{\sum_{j=1}^{N} \gamma_{jk}},\ \alpha_k = \frac{\sum_{j=1}^{N} \gamma_{jk}}{N};\ k = 1, 2, ..., K$

4.1.1. HYPERPARAMETER TUNING

The total number of Guassian components, K, is a hyperparameter that is given to the model. Thus, we need to conduct hyperparameter tuning. In order to conduct model selection, we use BIC to determine the optimal regime number. BIC is given by the formula $BIC = -2\, logL(\theta) + d\, log(N)$, where N is the sample size of the training set and d is the total number of parameters. BIC score gives an estimate of the model performance, lower BIC score signals a better model. We have applied our entire data set to the GMM with different number of regimes, K, ranging from 2 to 10. We found that $K = 4$ generates the lowest BIC. However, it should be noticed that:

- The optimal choice of 4 is based on using the entire data set. While during our backtesting, we are using a rolling window approach, where each rolling training set is just a section of the entire data set. So during the backtesting, instead of fixing $K = 4$ for each rolling window, we also experiment with changing the number of regimes dynamically based on each year's training data's BIC score. We will show the results in later sections.
- Here, we use BIC to choose the optimal number of regimes, while when we are backtesting our regime-switching investment strategies, we would be using portfolio performances in terms of P&L as a metric. It should be kept in mind that a low BIC score doesn't necessarily mean a high performance.

4.1.2. SENSITIVITY ANALYSIS

The results of GMM depend on the initialization (the mean and covariance matrix for each Gaussian component), which is the first step of the EM algorithm.To ensure $K = 4$ is the optimal choice for the number of regimes irrespective of initialization, we try several random initializations. For each choice of K, we run the model with 100 random initializations, record the BIC scores, and draw the box plot. As shown in Figure 6, the value $K = 4$ generates the lowest BIC irrespective of initialization.

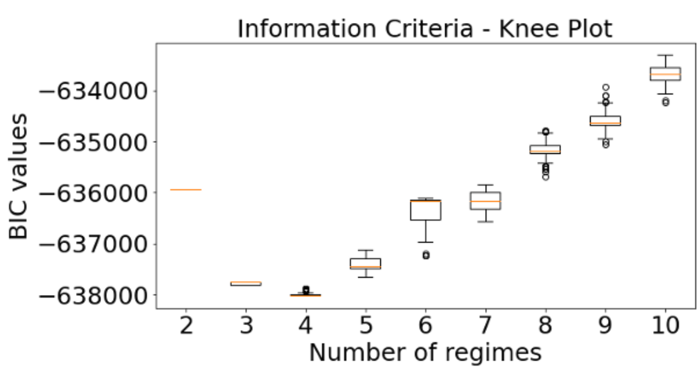

Figure 6: BIC Box-plots for different number of regimes

4.2. IN-SAMPLE REGIME DETECTION

Choosing the optimal hyperparameter and ensuring the mode is not sensitive to initialization, we now choose one of these initialisations and run the GMM model. The GMM model produces the probabilities of being in a pre-defined regime. The model selects the regime with the highest probability as the test day's regime. Applying the GMM model to the first 13 years of daily data, with fixed optimal 4 regimes as the BIC value suggests, the in-sample regime detection result is shown in Figure 7. In the plot, there are in total 4 colors, from light blue, yellow, navy to gray, representing 4 different regimes. The colors are assigned based on the total count of each regime during the 13 year time period in descending order. Therefore, the most common regime is assigned as light blue, while the rarest regime is assigned as gray.

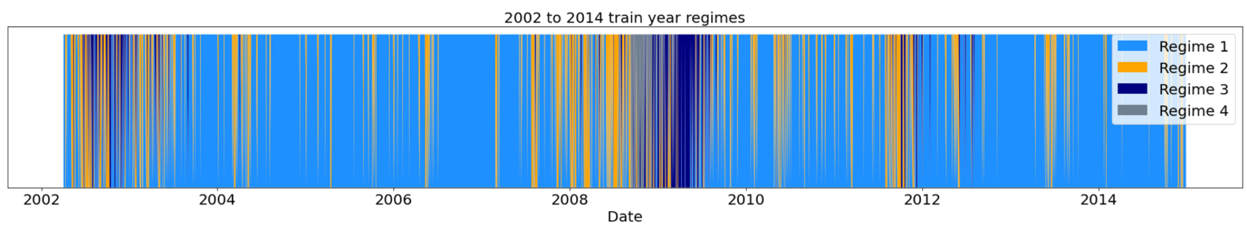

Figure 7: In-sample Regimes

4.2.1. SMOOTHING

The economic states are expected to last for a period of time, such as a bear or bull market. However, the regimes in Figure 7 have a high switching frequency. Our results do not match our economical intuition and may lead to high rebalancing costs when constructing a portfolio. Therefore, the results should be smoothed to a more robust and stable regime output. Three kinds of smoothing methods have been applied: probability threshold, probability average and regime maximum count.

1 - Probability Threshold Smoothing:

In the initial result, the model selects the regime with highest probability as the test day's regime on a daily basis, regardless of whether there is significant difference among the probabilities. Probability threshold smoothing changes the regime only if the regime with highest probability exceeds a certain threshold on the test day. Otherwise, the test day's regime will remain the same as the previous day. There are around 70% of samples that have the highest regime probability larger than the threshold at 90%, 95% and 99%. 30% of the regime switching frequency will be eliminated. However, the threshold needs to be manually selected as an additional parameter. It adds to the model complexity.

2 - Probability Average Smoothing

In the probability average smoothing method, the model takes previous 5 trading days' respective regime probabilities and calculates the average of the 5 days probabilities, both arithmetically and exponentially. The exponential average gives more weights on recent days than earlier days. The regime with the highest average probability will be selected as the test day's regime. The probability average smoothing method has a similar problem as the probability threshold method, which is the additional parameter introduced by the exponential weighted smoothing factor alpha. It essentially represents the weighting applied to the most recent period, but again adds to the model complexity.

3 - Regime Maximum Count Smoothing

The Regime Maximum Count smoothing method will use the GMM model to predict previous 5 days results and take the majority of the regime as the test day's regime. The smoothing method efficiently reduces the regime switching frequency. It could avoid the additional parameter as the previous 2 methods and stick to the portfolio's weekly rebalancing strategy.

Hence, regime maximum count smoothing is chosen as the optimal smoothing method in this project. The comparison with the unsmoothed results is shown in Figure 8. From the red circle in the plot, the short periods of regime 2 in yellow are largely eliminated while the other regime colors also last for a relatively longer timer period.

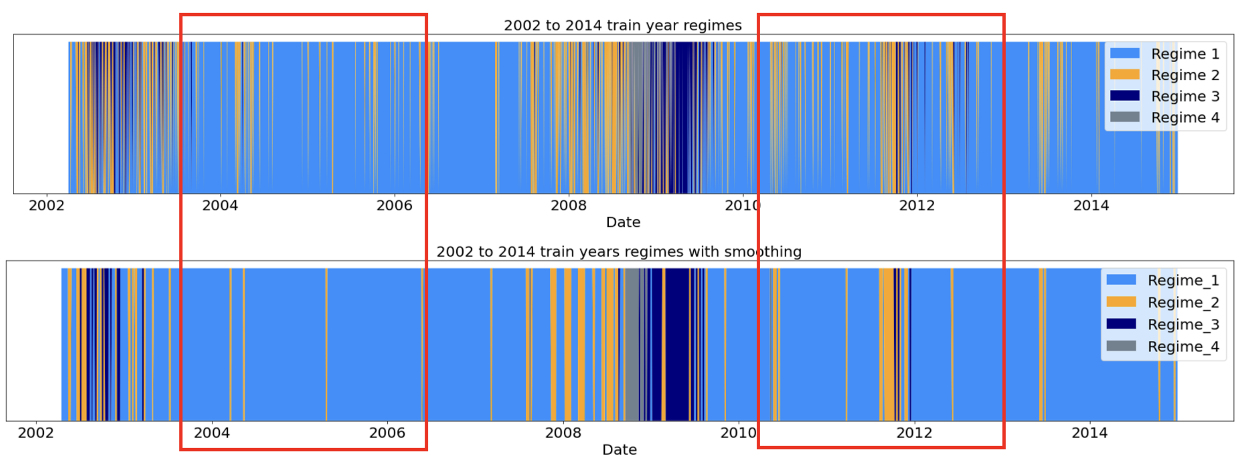

Figure 8: In-sample Regimes after Smoothing

4.2.2. REGIME INTERPRETATION

After the smoothing, the means and standard deviations of each regime are calculated and shown in Figure 9. The statistics of the regimes based on the first training window, from 2002 to 2014, are recorded to map the detected regimes to real world events.

	Regime 1 Mean	Regime 2 Mean	Regime 3 Mean	Regime 4 Mean	Regime 1 Std	Regime 2 Std	Regime 3 Std	Regime 4 Std
Interest Rate	0.060956	0.009458	0.073174	0.200154	0.048447	0.071136	0.071689	0.114690
Equity	0.262319	-1.114486	0.835393	-1.711932	0.097479	0.219833	0.281063	0.601352
Credit	0.112441	-0.054477	0.295093	-0.801853	0.021939	0.033447	0.066406	0.098004
Commodities	0.192342	-0.501196	1.076065	-3.378785	0.191519	0.289661	0.357263	0.559356
Emerging Market	0.255629	-0.754731	0.895947	-2.101479	0.072080	0.149534	0.152968	0.399401
Foreign Currency	0.017481	-0.109800	0.087853	-0.300355	0.020389	0.033859	0.040173	0.105594
Local Inflation	0.062209	0.078145	-0.002724	-0.187859	0.040287	0.056241	0.059197	0.121440
Equity Short Volatility	0.196107	-0.593692	0.517035	-1.072698	0.056880	0.203157	0.166216	0.563612
Fixed Income Carry	0.048167	0.046269	0.043584	-0.127010	0.044101	0.063856	0.050224	0.074331
Foreign Exchange Carry	0.164027	-0.530304	0.554713	-1.244218	0.062782	0.124929	0.122470	0.385784
Trend Following	0.100906	-0.165209	-0.147152	0.470582	0.039223	0.082200	0.066149	0.139277
Low Risk	0.127774	-0.216885	-0.055509	-0.417797	0.040025	0.063891	0.124430	0.184159
Momentum	0.131012	-0.118225	0.077648	-0.276589	0.060382	0.116637	0.189817	0.235941
Quality	0.024200	0.004327	0.048893	-0.007269	0.027160	0.034947	0.054194	0.055376
Value	0.021902	0.087450	0.396985	-0.189165	0.036219	0.050198	0.079704	0.088722

Calculation on train set of 2002-2014

Figure 9: Statistics of Factors in in-sample Regimes

	Regime 1	Regime 2	Regime 3	Regime 4
Observation from mean	Almost all factors have positive average returns	Most factors have negative returns, except Fixed Income, Inflation and Value	Many factors have high returns, especially Equity and Commodity	Almost all factors have negative average returns
Observation from variance	Low volatility in all factor returns	High volatility in Commodity	Higher volatility than regime 1	High volatility for all factor returns
Regime label	Steady State	High Inflation	Walking on Ice	Crisis

Table 4: Regime Interpretation

Here, Regime 1 displays all positive returns for all 15 factors. The standard deviations of Regime 1 are low. Regime 1 is the most common regime. Therefore, Regime 1 likely represents the Steady State. On the contrary, Regime 4 has almost all negative returns for the 15 factors and high volatility. Therefore, Regime 4 is referred to as the Crisis State. For Regime 2, while most factors show a negative return, FI, Inflation and Value Factors have a positive mean. Commodity has high volatility. Given the characteristics of different factors and their performance, Regime 2 may indicate the High Inflation State. Equity and Commodity factors demonstrate strong return in Regime 3, however, they also come with high volatility. Therefore, Regime 3 is referred to as the Walking on Ice State. These characteristics and regime labels are shown in Table 4.

Applying the regime labels to the real world scenario, as shown in Figure 8, most of the time are in the Steady State (Regime 1 in light blue). During the 2008 Global Financial Crisis, the detection result firstly showed High Inflation States (Regime 2 in yellow). Then, it is followed by the Crisis State (Regime 4 in gray) for a short period of time. Walking on Ice State (Regime 3 in navy) came after it and lasted for a relatively long and consecutive time. The mapping to history proves that the regime detection result is reliable and the model is robust.

5. PORTFOLIO CONSTRUCTION

Since we already have a basic understanding of the economic intuition behind each regime, we try to figure out how we can leverage this information to build effective investment strategies. Here, we focus on two kinds of portfolios: Top Sharpe portfolio and Markowitz portfolio.

5.1. TOP SHARPE PORTFOLIO

Across different regimes, the dominance of the factors would be drastically different. Therefore, we can pick up the factors with best performance in each regime and focus on those factors when building the portfolios. To be specific, once we fit the GMM model, we can get a bunch of mean vectors and covariance matrices, which could further yield the sharpe ratio for each factor after simple calculation. Then the portfolio for each regime would be simply constructed by the two highest sharpe ratio factors correspondingly. They would be equally weighted in the portfolio. The Table 5 below summarizes our factor selection for each regime.

	Steady State	High Inflation	Walking on Ice	Crisis
Factor 1	Credit	Value	Emerging Market	Trend Following
Factor 2	Emerging Market	Local Inflation	Value	Interest Rate

Table 5: Top Sharpe Portfolio Factors

5.2. MARKOWITZ PORTFOLIO

Classical Markowitz portfolio optimization for long only portfolio solves the following optimization problem to allocate weights to companies within the portfolio, where γ is the risk aversion parameter and the optimal weights selected are corresponding to the γ that gives the highest Sharpe Ratio:

$$max\ \mu^T w - \gamma w^T \Sigma w \quad st.\ 1^T w = 1,\ w > 0,\ w \in W$$

In Figure 9, we plot the efficient frontier along with three different types of portfolio for each regime. The circle represents the highest sharpe ratio portfolio. The plus sign is the minimum variance portfolio, and the triangle is the inverse volatility portfolio. We can notice that the inverse volatility portfolio is always inside the efficient frontier and the minimum variance portfolio is almost always the leftmost bottom point on the curves since the curves are not exactly parabolas.

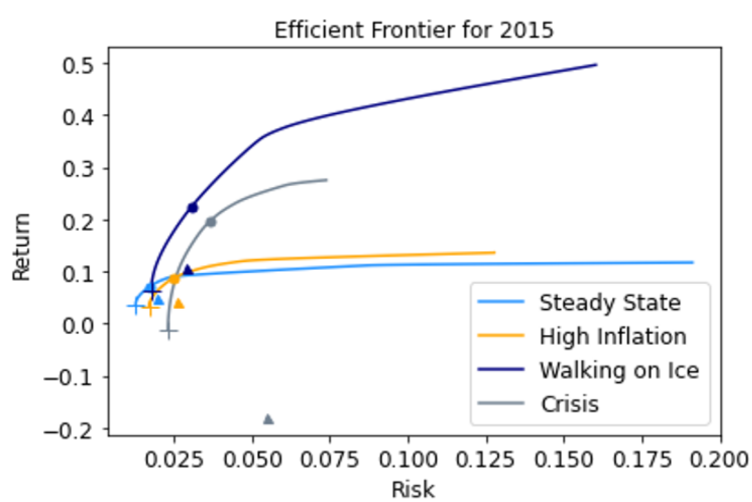

Figure 9: Efficient Frontiers

	Steady State	High Inflation	Walking on Ice	Crisis
Interest Rate	0	0	0	0.4328
Equity	0	0	0	0
Credit	0.4862	0	0.1628	0
Commodity	0.0041	0	0	0
Emerging Market	0.0173	0	0.1321	0
Foreign Currency	0	0	0	0
Local Inflation	0	0.3768	0	0
Equity Short Volatility	0.1217	0	0.0372	0
Fixed Income Carry	0.0506	0.0910	0.3205	0
Foreign Exchange Carry	0.0387	0	0.0540	0
Trend Following	0.0363	0	0.0425	0.5671
Low Risk	0.0597	0	0.0588	0
Momentum	0.0506	0	0	0
Quality	0.0580	0	0.0945	0
Value	0.0761	0.5321	0.0973	0

Table 6: Markowitz Portfolio Weights

In Table 6, we list the Markowitz portfolio weights of the highest sharpe ratio in each regime. Some of the factor weight makes a lot of sense. For example high inflation regime portfolio corresponding to a high position of Local Inflation factor. In Steady State, we will tend to hold a more diversified portfolio while in Crisis, we tend to hold a less diversified portfolio. We hold Local Inflation during the High Inflation regime and Equity Style factors in Walking on Ice. All these portfolio weights seem to match what we intuitively expect to happen in these market regimes. One issue that we may feel concerned is when the regime switches, the position of the portfolio would be drastically changed. A possible solution is instead of just holding the optimal Markowitz portfolio, we could always hold a combination of all the regime portfolios at the same time, weighted by their corresponding probabilities of regimes given by GMM. Further details on this will be discussed in later sections. Another alarming issue here seems to be that Equity does not get any weight in any regimes, especially in the Steady State regime where it was expected. Instead Credit seems to receive weight then. Further analysis is conducted on this issue to see if the weights allocation can be explained intuitively from the model as well from the economic sense from the historical time period covered in the training data. Details on this are provided in the following subsection.

5.2.1. EQUITY VS. CREDIT

Here we explore more about whether or not zero weights on equity is a concern. Technically, it is not that equity gets no weight at all because in our investment factors, we have some equity style factors including momentum, quality, value etc. which have a non-zero weight. So in some regimes, we do have exposure to the equity market.

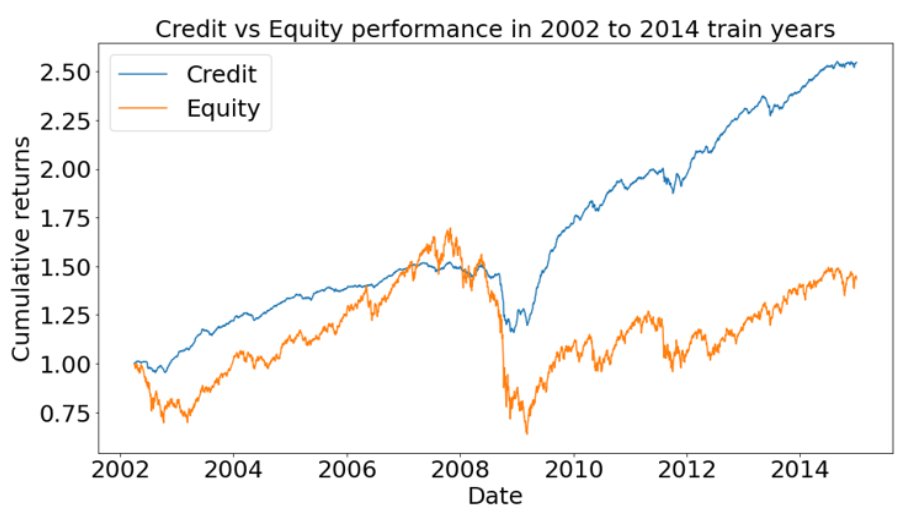

Figure 10: Cumulative Returns for Equity vs. Credit

Another thing to note here is that the Credit factor seems to be more dominating over the Equity factor for all the regimes. This can be intuitively explained as follows :

- Intuition from Economics: All the training periods include post 2008 financial crisis period in which Credit was an attractive investment because of the Fed's credit easing policies. Meanwhile, the financial crisis resulted in a very bad performance of the equity market. This can be seen in Figure 10.
- Intuition from Model: Markowitz focuses on risk adjusted returns, which is better for Credit than Equity as we can see from Table 7.

	Annualized Returns (%)	**Annualized Volatility (%)**	**Sharpe Ratio**
Equity	2.80	17.53	0.25
Credit	7.45	3.45	2.19

Table 7: Statistics for Equity vs. Credit

6. RESULTS

In the following subsections we present the methodology and result of out of sample regime detection and comparison of no regime vs. regime based portfolio performance.

6.1. OUT OF SAMPLE REGIME DETECTION AND PORTFOLIO CONSTRUCTION

In order to identify regimes out-of-sample, we use a rolling-window approach, with test years ranging from 2015 to 2022. We roll our training and test data set yearly as shown in Figure 11, using past 13 years training data to recalibrate the model for every test year.

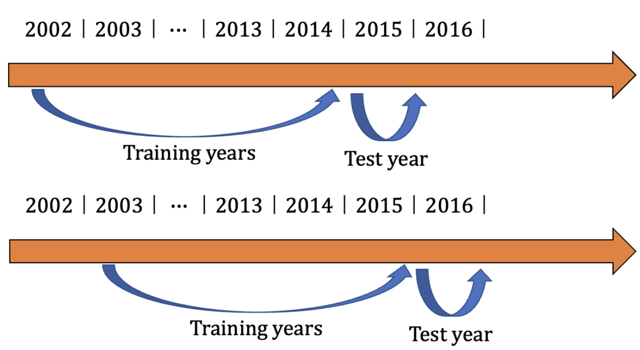

Figure 11: Rolling Window Framework

We classify regimes in test year using the GMM model trained on training years. The out-of-sample regimes after smoothing are shown in Figure 12 below.

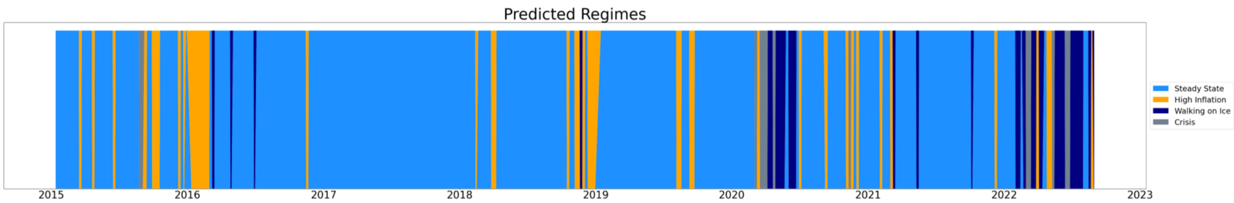

Figure 12: Out-of-sample Regimes

From Figure 12, we can see that our model identified the Covid crisis and current inflation market ups and downs. So the model does seem to work robustly in all market conditions. Following Figure 13 shows changing portfolio weights in these out-of-sample regimes based on regime-wise optimal portfolios constructed before.

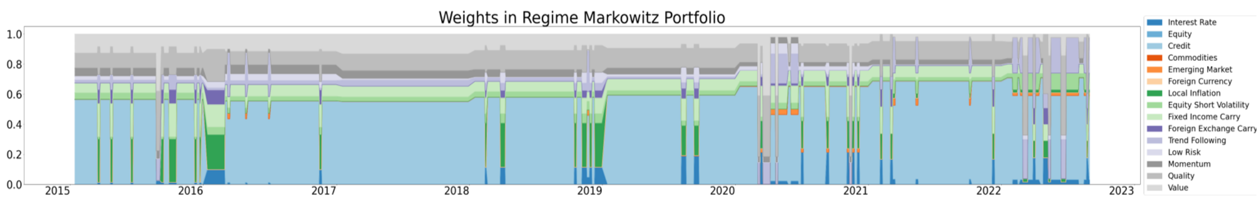

Figure 13: Out-of-sample Portfolio weights

6.2. PERFORMANCE EVALUATION

Here we look at the summary of the performance of the two kinds of portfolios that we discussed, the Top Sharpe portfolio and the Markowitz portfolio, both constructed without and with the regime detection model, referred to as no regime portfolio and regime portfolio respectively. When we say a no regime portfolio, we just look at the entire train period and build a single portfolio from it that we hold in the test period. When we say a regime-based portfolio, we divide the train period into different regimes and make portfolios specific to each regime. Then we predict regimes in the test period and hold portfolios according to that. In the following subsection we look at performance metrics that we plan to use for evaluating and comparing all the portfolios.

6.2.1. EVALUATION METRICS

The performance metrics used here are as described in Table 8 below.

	Description
Return	Compounded annual growth rate of daily returns
Sharpe Ratio	Annualized risk adjusted returns
Max Drawdown	Maximum decline from previous peak
Calmar Ratio	Annualized returns given maximum drawdowns

Table 8: Performance Metrics

6.2.2. PORTFOLIO COMPARISON

Having described the performance metrics for portfolio comparison, following Table 9 shows how each portfolio does with respect to each of these. Later, Figure 14 shows the performance of different portfolios during the test period. We compare all our portfolios with a simple equal weighted portfolio in all assets during the test period.

	Annualized Returns (%)	Sharpe Ratio	Maximum Drawdown (%)	Calmar Ratio
No Regime Top Sharpe	-0.91	-0.21	15.00	-0.05
Regime Top Sharpe	0.41	0.10	21.17	0.03
No Regime Markowitz	0.27	0.12	11.13	0.03
Regime Markowitz	2.30	0.95	5.42	0.43
Equal Weighted	1.94	0.46	13.94	0.14

Table 9: Statistics for Model Comparison

A very powerful observation here is that regime detection always seems to add value to the regime-based portfolio as compared to the no regime portfolios and this true for both Top Sharpe and Markowitz. We also see that Markowitz does better than Top Sharpe, possibly because of two reasons, first it allows flexibility to assign weights to all assets and second it takes into account the correlation between different assets too. One final important thing to note here is that it may look surprising as to why an equal weighted portfolio is doing better than optimized portfolios at times. A possible explanation for this is the fact that optimal portfolios are constructed based on the in-sample data, but the distribution of out-of-sample data might differ, and the constructed optimal portfolio may not have the best performance out there. While we have no specific way to ensure these distributions are the same and can only hope so, we will look into potential improvement of portfolio construction methodology towards the end. But overall based on all the metrics we conclude that Markowitz portfolio constructed using regime detection results is the one to go with.

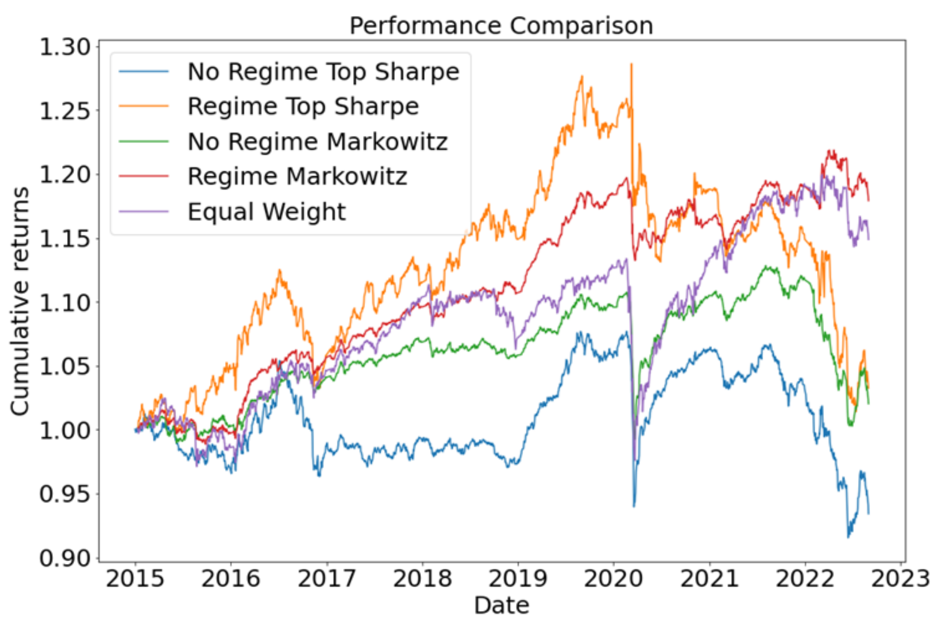

Figure 14: Cumulative Returns for Model Comparison

7. EXPERIMENTS

Having narrowed down to the best portfolio construction method, we now conduct several experiments to improve performance. Each of these experiments make a slight modification to the original model settings of regime based Markowitz portfolio and analyze the performance change. Following Table 10 gives a summary of the results of all the experiments. Later subsections provide more details.

	Annualized Returns (%)	Sharpe Ratio	Maximum Drawdown (%)	Calmar Ratio
Dynamic Number of Regimes	2.27	0.99	5.30	0.43
Weekly Data for Regime Detection	0.29	0.13	8.22	0.04
Portfolios Weighted using Probabilities	1.61	0.76	4.49	0.36
Expanding Window Training	2.92	1.21	6.21	0.47
Equity Weights Boosting	3.23	0.76	10.55	0.31
Model Generalization for US Investors	5.35	1.30	9.9	0.53

Table 10: Experiment Result Summary

7.1. DYNAMIC NUMBER OF REGIMES

We change the optimal number of regimes dynamically. Rather than sticking with 4 regimes throughout, every time we roll the training window to recalibrate the model, we calculate the optimal number of regimes based on BIC score in the new training window and build portfolios accordingly to be used in the test period.

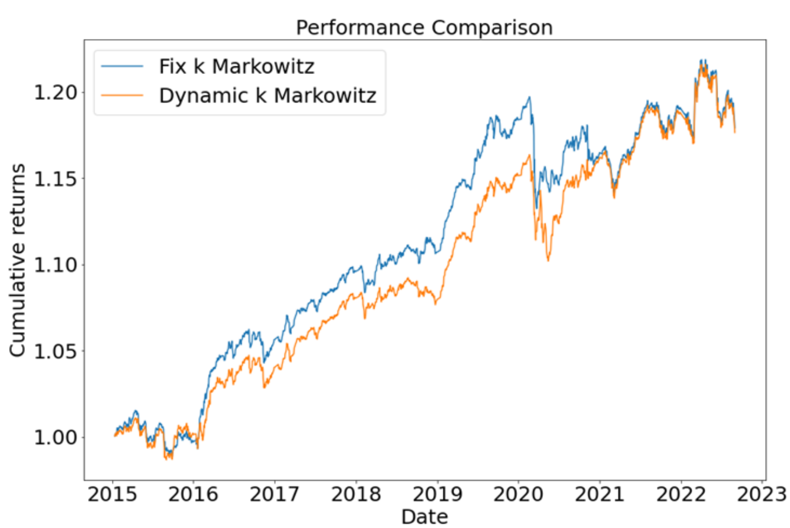

Figure 15: Fix vs. Dynamic number of regimes

We see it has comparable results with the original model but by changing the number of regimes, we have to map the detected regimes to real world regimes at every recalibration. Also, we choose the optimal number of regimes dynamically based on BIC score but we evaluate model performance based on PnL. The final model may not have the best out of sample performance. We cannot choose based on PnL because it is an incremental metric i.e., with more regimes we capture more granular trends and get better PnL so PnL would always suggest the highest number of regimes and not penalize the model complexity that comes with it. Following Figure 15 compares the performance of original regime based Markowitz portfolio with a fixed 4 number of regimes with regime based Markowitz portfolio with number of regimes changing dynamically.

7.2. WEEKLY DATA FOR REGIME DETECTION

We use a weekly average data instead of daily data for regime detection. So instead of detecting daily regimes and smoothing it to convert into weekly, we directly use weekly data and detect weekly regimes. However, it looks like this approach does not work well because the model has less information to use to identify regimes accurately. Following Figure 16 compares performance of original regime based Markowitz portfolio based on daily regime detection with regime based Markowitz portfolio with weekly regime detection.

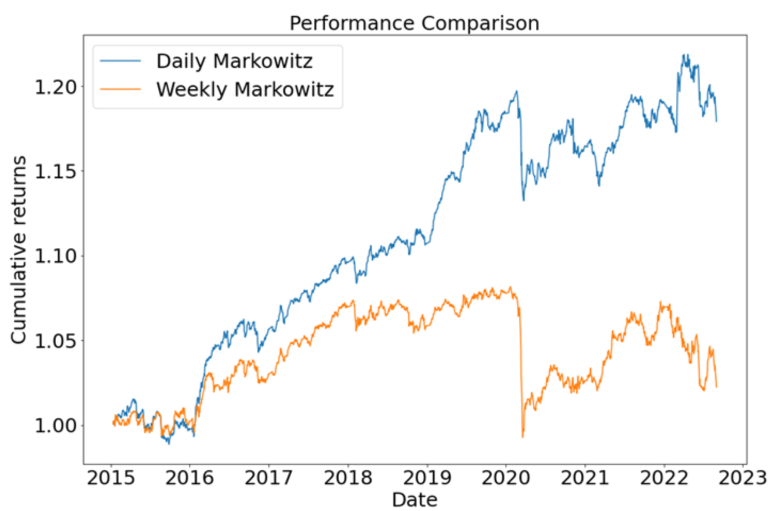

Figure 16: Daily vs. Weekly data for Regime Detection

7.3. PORTFOLIOS WEIGHTED USING PROBABILITIES

Here, we weigh portfolios using regime probabilities. We know that the GMM model gives probabilities of being in all regimes and we pick the regime with maximum probability and hold the portfolio for that regime. Now instead of that, we hold portfolios of all regimes, but weighted by their corresponding probabilities given by GMM. This will help us save on transaction costs by not having to switch between two completely different portfolios when regimes change, but just adjust the holdings slightly with every switch. But the trade-off here is that we get poorer returns and Sharpe because we may be holding assets not recommended for high returns specifically for this regime, but we have lower drawdowns because of the diversification. Following Figure 17 shows how weights of different assets held during the test period changes and we clearly see that with weighted portfolios the positions do not change drastically as before.

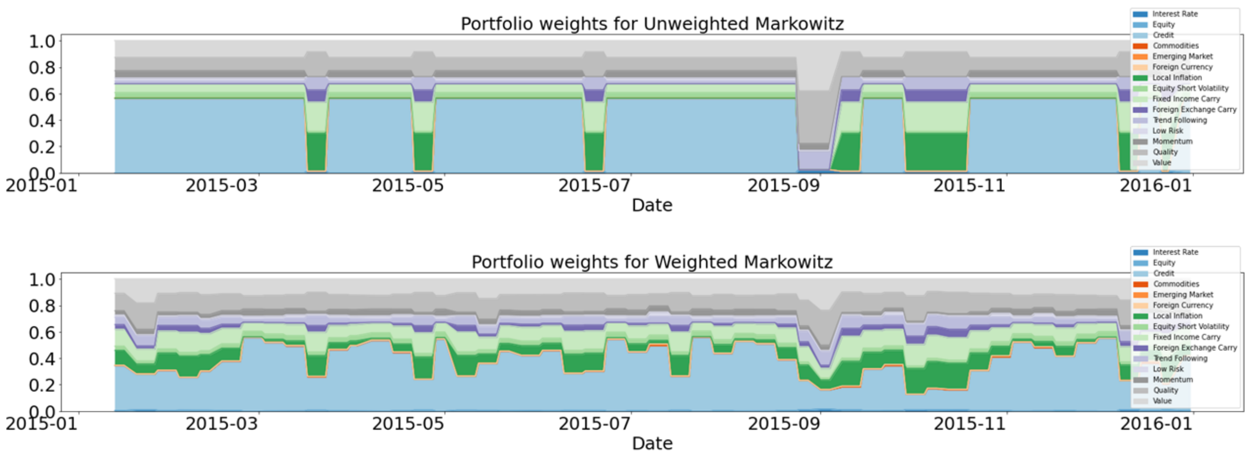

Figure 17: Unweighted vs Weighted Markowitz Portfolio

7.4. EQUITY WEIGHTS BOOSTING

Here, we put a minimum weight on equity. Although we justified why equity did not have an explicit weight, there might be investors who still want equities in their portfolio, so we keep minimum 20% weights on equities and then optimize for all assets.

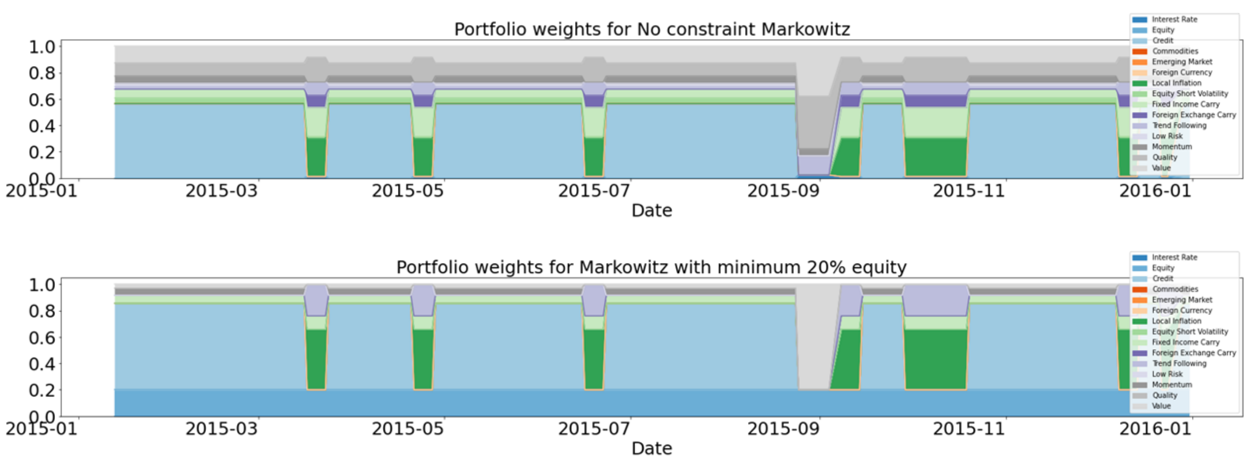

Figure 18: No constraint vs. constraint on Equity weight in Markowitz Portfolio

We see that as expected, with more equity, we get more returns, but sharpe drops because of more risk. Following Figure 18 shows the weights of different assets held during the test period with and without the minimum equity constraint. We see that with constraint, equity is given close to the minimum weight forced by the constraint and nothing more.

7.5. EXPANDING WINDOW TRAINING

Here, we use an expanding window for training. So rather than using just past 13 years for training, we use all years starting from 2002 until the test year as the training data. With this we get better results because the model has more data to learn from. Figure 19 compares performance of regime based Markowitz portfolio with rolling window training with regime based Markowitz portfolio with expanding window training.

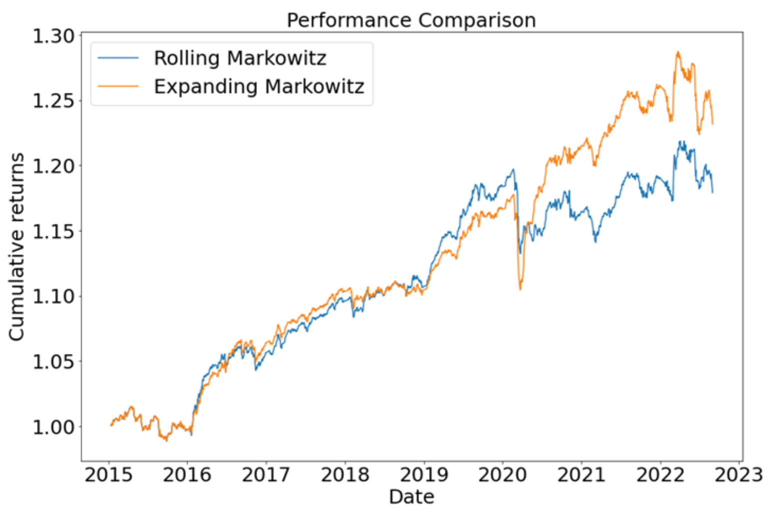

Figure 19: Rolling vs. Expanding Window Training for Regime Detection

7.6. MODEL GENERALIZATION

Finally, here we use a totally new dataset with US specific factors for US investors instead of the global factors and see if the model can generalize well over that too. Following Table 11 describes the details of the factors used for this US specific portfolio. We observe quite an improvement in performance of the US portfolio as compared to the global portfolio, mainly because the equity factor used here i.e. S&P 500 has done much better historically than the equity factor MSCI World used in the global portfolio. so it makes sense for US investors to hold this portfolio instead, if they are willing to bear a slightly higher drawdown. Figure 20 compares the original regime based Markowitz portfolio on global factors with the regime based Markowitz portfolio on US specific factors.

	Factor	Construction Details	Bloomberg Tickers
Core Macro	Equity	S&P 500 Index	SPX
	Interest Rates	US Treasury Yield 10Y	H15T10Y
	Credit	Average of Bloomberg Barclays US Corporate High Yield (US HY), and Bloomberg Barclays US Corporate (US IG)	LF98TRUU LUACTRUU
	Commodities	S&P GSCI index, a benchmark for commodity markets	SPGSCI
Secondary Macro	Emerging Markets	Average of two indices: MSCI Emerging markets Index (EM equity), and Bloomberg Barclays Emerging Markets USD Aggregate (EM Credit)	MXEF EMUSTRUU
	Local Inflation	Bloomberg Barclays US Treasury Inflation-Linked Notes	LBUTTRUU

Table 11: Factor Construction for US specific factors

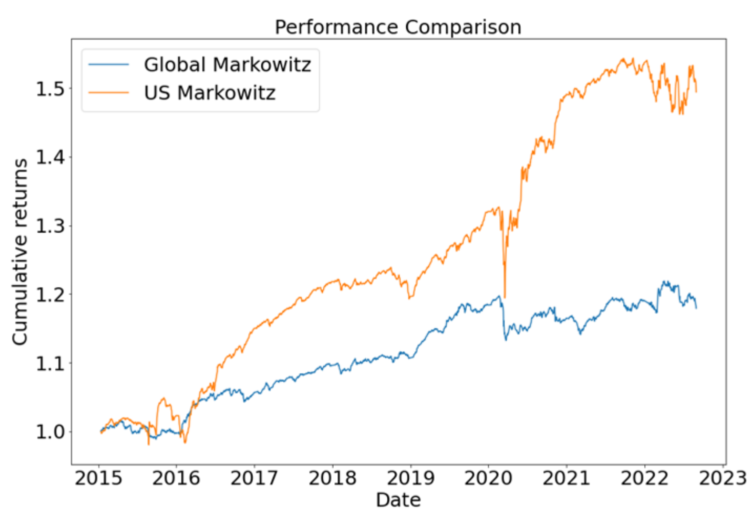

Figure 20: Performance of Global vs US specific Markowitz Portfolio

8. CONCLUSION

Our study explores the regimes or clusters in the financial market with specific properties. Each regime had different properties and therefore each asset performed differently in these regimes. One approach of identifying such regimes is with human observation and market knowledge while the other approach is data driven with various factors. The study explored the data driven approach with the use of an unsupervised learning algorithm in the form of Gaussian Mixture Model. With GMM, we detected clusters or regimes in the market using the dataset mentioned in the earlier section. It is important to note that models like GMM are sensitive to the hyperparameters and these need to be tuned as we have done in this study. Along with detecting the market regimes, we have labeled them based on their market conditions using the mean returns and standard deviations as Steady, high Inflation, Crisis and Walking on Ice states. We analyzed the behavior of these regimes and saw persistence in their properties throughout history.

For the use case of market regime detection, the study explored portfolio construction for asset allocation. Even in extremely stable market conditions, predicting market returns over a lengthy period of time is challenging. In terms of asset allocation, this means that one should create portfolios that can endure significant volatility and drawdowns over the long run while also perhaps looking for opportunities in the shorter term. Therefore, we examined the performance of the Markowitz portfolio, a conventional mean variance configuration, both with and without regime detection. As per our results that we saw in the previous section, the regime detection model can capture superior performance of certain assets in each regime. As a result, this improves the performance of the entire portfolio. Additionally, we observe that regime detection often results in fewer drawdowns, which contributes to overall superior compounding returns.

The project also resulted in streamlining the code to make it more adaptable to parameter changes. For instance, altering the universe of factors, having an expanding window, or dynamically varying the number of regimes throughout rolling windows, among other examples. The pipeline is designed to be more of a plug-and-play code, that enables one to simply alter the parameters for future studies. The github link to the project can be found at the end of this document.

Following is a brief discussion on the future scope of this research.

In this study, we used Gaussian Mixture Model to detect market regimes. The Hidden

Markov Model is another widely used way to detect regimes in the industry. With the Hidden Markov model, we use the observations(returns) to determine the hidden states that drive these observations. We can interpret these hidden states as Regimes in the market. Therefore, HMM can add value to the pipeline presented in this study with its distinctive way of identifying the regimes.

The study also saw that Markowitz portfolio weights are high for Credit in our labeled Steady state. This is in contradiction to the intuition that one may want to have more exposure to Equity when the market is steady. On further analysis, we saw that the training set for GMM has more data points where Credit performed better than Equity and we also presented the economic intuition for these data points. But the more recent data in the Steady states suggest that indeed Equity is the better investment. To tackle this issue, we can have some variation in the model. First, we can experiment with tweaking the GMM model to prioritize current data so that the model is savvy enough to recognize recent patterns in addition to general trends to identify regimes. Another option is to build the portfolio over a relatively short period of time while training the model over a longer period of time to build a portfolio capturing recent developments.

BIBLIOGRAPHY

1. Geoff Duncombe and Bradley Kay (2018). "Introducing the Two Sigma Factor Lens"
2. Alex Botte and Doris Bao (2021). "A Machine Learning Approach to Regime Modeling"

*Please find data/code here - https://github.com/DN612/Capstone-Regime-Detection